



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 164752

TO: Phuong Bui
Location: rem/2A15/2C18
Art Unit: 1638
Saturday, September 10, 2005

Case Serial Number: 10/829432

From: Mary Jane Ruhl
Location: Biotech-Chem Library
Remsen 1-A-62
Phone: 571-272-2524

maryjane.ruhl@uspto.gov

Search Notes

Examiner Bui,

Here are the results for your recent search request.

Please feel free to contact me if you have any questions about these results.

Thank you for using STIC services. We appreciate the opportunity to serve you.

Sincerely,

Mary Jane Ruhl
Technical Information Specialist
STIC
Remsen 1-A-62
Ext. 22524



THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 9, 2005, 09:19:19 ; Search time 5327 Seconds
(without alignments)
11070.020 Million cell updates/sec

Title: US-10-829-432-3
Perfect score: 1217
Sequence: 1 ggcgtcgcttcatcattca.....aaaaaaaaaaaaaaaaaaaa 1217

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_com.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ste.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	359.6	29.5	928	8	BT009518
2	333.2	27.4	1434	8	AK071285 Oryza sat
3	311	25.6	970	8	AF178976 Zea mays
4	309	25.4	1086	8	AF044285 Atharant
5	296.4	24.4	1068	8	AY085264 Arabidops
6	296.4	24.4	1077	8	AY05238 Arabidops
7	296.4	24.4	1143	8	AY054287 Arabidops
8	296.4	24.4	1185	8	ATC14PSK
9	294.4	24.2	831	6	CO805858 Sequence
10	294.4	24.2	831	6	AX412633 Sequence
11	294.4	24.2	831	6	AX506015 Sequence
12	294.4	24.2	831	6	AY132010 Arabidops
13	288.4	23.7	964	8	BT005193 Arabidops
14	288.4	23.7	1133	8	AY085031 Arabidops
15	288.4	23.7	1143	8	BT003977 Arabidops
16	263.6	21.7	882	8	AY097421 Arabidops
17	263.6	21.7	920	6	BD248389 Gene part
18	263.6	21.7	1243	8	AF462823 Arabidops
19	263.6	21.7	1311	8	AF043351 Arabidops

20	172.6	14.2	300150	1	AP004598	AP004598 Oceanobac
21	170.2	14.0	11131	1	AE007523	AE007523 Clostridi
22	168.6	13.9	11037	1	AE013685	AE013685 Versinia
23	168.6	13.9	110000	1	EX936398_09	Continuation (10 o
24	168.6	13.9	220050	1	AJ414156	AJ414156 Versinia
25	168.6	13.9	290803	1	AE017128	AE017128 Versinia
26	165.2	13.6	349742	1	EX572090	EX572090 Prochloro
27	163.4	13.4	51860	8	AB013390	AB013390 Arabidops
28	162.8	13.4	521	8	BT009596	BT009596 Triticum
29	161.4	13.3	53533	1	BSY09476	Y09476 Bacillus su
30	161.4	13.3	198743	1	BSUB0006	Z99109 Bacillus su
31	160.2	13.2	300275	1	AE016751	AE016751 Staphyloc
32	159.6	13.1	302050	1	AL935256	AL935256 Lactobaci
33	159.2	13.1	349814	1	CR378673	CR378673 Photobact
34	156.8	12.9	153950	8	AC104429	AC104429 Oryza sat
35	156.4	12.9	609	6	CO871473	CO871473 Sequence
36	156.4	12.9	1148	8	AY017216	AY017216 Saccharom
37	156.4	12.9	1160	6	AK179484	AK179484 Sequence
38	156.2	12.8	300000	1	AP005073	AP005073 Vibrio pa
39	155.6	12.8	2372	3	AK173431	AK173431 Ciona int
40	155.6	12.8	2378	3	AB036852	AB036852 Ciona int
41	155.6	12.8	2380	3	AK173392	AK173392 Ciona int
42	155.2	12.8	1848	4	AY618932	AY618932 Oryctolag
43	155.2	12.8	147205	8	AC146702	AC146702 Genomic s
44	154.8	12.7	2080	8	AB049836	AB049836 Saccharom
45	153.8	12.6	1021	11	CNS06JUF	AL401665 T7 end of

ALIGNMENTS

RESULT 1
BT009518
LOCUS BT009518 928 bp mRNA linear PLN 20-JUN-2003
DEFINITION Triticum aestivum clone wrl.pk0101.e2:figs, full insert mRNA
ACCESSION BT009518
VERSION BT009518.1 GI:32129069
KEYWORDS FLI CDNA.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 928)
AUTHORS Tingey,S.V., Wolters,P., Powell,W., Dolan,M., Miao,G.-H.,
Caraher,N.R., Hanafey,M.K. and Hainey,C.F.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-2003) Crop Genetics, E. I. Dupont de Nemours and
Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,
USA

FEATURES

source Location/Qualifiers
1..928
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wrl.pk0101.e2:figs"

ORIGIN

Query Match 29.5%; Score 359.6; DB 8; Length 928;
Best Local Similarity 65.5%; Pred. No. 1.6e-76;
Matches 558; Conservative 0; Mismatches 289; Indels 5; Gaps 2;
QY 367 GCCCGGGAGCCCGCAGCCAGTGAAGAGAGCTGTAAATGTCGAACATTTGGGA 426
DB 77 GCCCGTCAATGGATCAGCCATGGCAGGTATCGAACAGTTGTGACCTCAACTGTTGGGAA 136
QY 427 ATCGACTAATATTTTATGGCAATTTGCTTGTATGGCAATCTGATAGACAGAAATTTGCT 486
DB 137 ATCGACAAACGTTCTTTGGCATGACTGTCCATAGTACTCAGTTTGAGAGCGGAGAACTGCT 196
QY 487 GGGACAAAAGCTGTCGTATGGATCAACAGGACTCAGTGGTTCAGGGAAAAGTACTCT 546

Db 197 AAATCAGAAGGGTTGTGTGTGGATAACAGGGTTAAGTGGTTTCAGGGAAAAAGCACACT 256
QY 547 TGCATGTGCACGAGTCGTCGAGTTCGATTCGAGAGCCACCTCAGTATGATCTTATGATGG 606
Db 257 AGCATGCGCGCTAAGTCGAGCTGCACCTCCAGAGGTCTATCGACTTACATTCCTAGACGG 316
QY 607 TGACAACTTCAGACATGGCTTAAATAGAGATTAAAGCTTTAAGGCAGAGAACCGTGCAGA 666
Db 317 TGACAACTTAAGCATGGGTAAACCGAGACCTCTGTTTCGAAGCAAGGACCGTGCCTGA 376
QY 667 AAATATACGAAGTGTGTGAAGTGCAAGAGCTTTTTCGTCGATGCTGTGTCATATGAT 726
Db 377 AAATATACGAGATGAGGAAGTAGCAAGAGCTGTTTCGAGATGCTGTGTCATCTGCAT 436
QY 727 TCGTAGCTTGATATCCATCAGAGAGATCGTGATGATCGCTGCTCTACTTCCACA 786
Db 437 TCGTAGCTTGATATCACCCTCAGAGAGTGAACGACGCTTGCCGCAATTTACTGCACAA 496
QY 787 TTCTAACTTTATTGAAGTATTTATTGATTTGCCCCCTTAAAAATTTGTGAAGCTTCGTGATCC 846
Db 497 TTCTACATTCATCGAGTGTGTTTGAATGTCCTCCACTTGAAGTTTGTGAAGCTTAGGATCC 556
QY 847 TAAAGCCTTATCAAGCTTGCACGTACAGGAAGATTAAAGTTTCACTGGAATTTGATGA 906
Db 557 AAAAGGCTTGTACAAGCTTGCCCGTGCAGAAAAAATCAAAGGGTTTACTGGAAATTTGATGA 616
QY 907 TCCATACGGAACCAATTAATGCTGAGATAGTAAATTAAGATGAAGATGAGGATGCC 966
Db 617 TCCATTATGAAGCACCTTCTGATCGGAGATAGTATACAGTGCNAAGCTGGTACTGCGC 676
QY 967 TTCACCCAAAGCAATGGCCCAAGCTTCTATGCTACCTTGAAGAAAAACGATATTGCA 1026
Db 677 CAGCGCTAATCGATGCTGATCAAGTTGTGTCATATCTTGAAGCAATGAGTCTTACA 736
QY 1027 AGCTTAG---TATATGATTTTGAGAAGATGATCTGATCTTGTGTCGATTAATCTTGT 1083
Db 737 GGAATGAGACGTATGCTATGATGAATAAATCTGAAATTTGGATCGCCAAAGGATGT 796
QY 1084 GCACACATAAGATCTGTTGTTGTCACATGAATAA--AAGCATCAACATCTAGGAAGT 1141
Db 797 GAAATATGAGGTAGTATTTATCTCTAGAAGAGTATGATAGTATGAGAACATATATAT 856
QY 1142 AACAGAAGTACGGTTTCATTCAGAAACGGATATGGATTTCATTCGTTTAAAAAATAAAAA 1201
Db 857 GACATAAAGATCGAATCTGATCATATATATTAATTTGAATGTTTTCAGCCCAAAAA 916
QY 1202 AAAAAAATAAAAA 1213
Db 917 AAAAAAATAAAAA 928

RESULT 2
AKO71285
LOCUS AKO71285 1434 bp mRNA linear PLN 24-JUL-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023086D14, full insert sequence.
ACCESSION AKO71285
VERSION AKO71285.1 GI:32981308
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team;
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Nami, T., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Ohtsuki, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,

Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN; Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Iishi, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
Science 301 (5631), 376-379 (2003)
22752273
12869764
2 (bases 1 to 1434)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, I., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kurosaki, T., Kusumegi, T., Kondo, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Matsuyama, T., Miura, J., Miyazaki, A., Masuda, H., Matsubara, K., Murata, M., Nagata, T., Nakamura, M., Nami, K., Nariawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami, T., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Nami, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Ohtsuki, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Iishi, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, C., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sugabe, Y., Tagami, M., Tagami, T., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.
Location/Qualifiers
1..1434
/organism="Oryza sativa (japonica cultivar-group)"

REFERENCE	2 (bases 1 to 970)																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
-----------	--------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

2

Malibu, CA 90265, USA

This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the ws or laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genet carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES

source

1. .1068
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="14216"
94. .924
/codon_start=1
/product="putative adenosine phosphosulfate kinase"
/protein_id="AA062496.1"
/db_xref="GI:21553403"

CDS

24.4%; Score 296.4; DB 8; Length 1068;
Best Local Similarity 68.0%; Pred. No. 4.1e-61;
Matches 430; Conservative 0; Mismatches 196; Indels 6; Gaps 1;
411 TGTGAAACATTGGGAAATCGACTAATATTTATGGACAAATTCCTTGGATGGACAACTCG 470
308 TGTCTACGGTCGGAACCTCGACAAATATAAAGTGGCATGAATGTCTCTGTGAGAAAGTTG 367
471 ATAGACAGAAATGTCTGGACAAAAGGCTGTCTGTATGGATACAGCACTCACTGGTT 530
368 ATAGACAGAGATGTCTTGATCAGAAGGATGTGTGATTTGGGTCAACGGTCTTAGTGGTT 427
531 CAGGGAAGAGTACTCTTGCATGTGCACCTGAGTGTGATTCAGAGGCCACCTCA 590
428 CAGGGAAGAGTACTTTGGCTTGTGCTTGAATCAGATGTTGATCAAAAGGGAGCTTT 487
591 COTATGTACTTGAATGGTGACACCTCAGACATGGCTTAATAGAGATTTAAGCTTTAAGG 650
488 GTTATATTCTTGAATGGTGATATGTTAGGCATGGCTTAAACCGTATCTTAGCTTTAAAG 547
651 CAGAGACCGTCGAGAAATATACGAGAGTTGGTGAAGTGGCAAGCTTTTGTGCTATG 710
548 CTGAGGATCGTCGAGAGATATTCGTAGAGTTGGAGAGTTGCTTAAGCTTTTTCGGATG 607
711 CTGGTGTCTATGCAATGCTAGCTGATATCTCCATACAGGAGAGATCGTGTATGATGCC 770
608 CTGGAATATCTGCATTTGGAGTTTGAATCTCTTATAGACATAGGAGCGTTGTC 667
771 GTGCTCTACTTCCATCTTAATCTTATTTAAGAGTATTTATGATTTGCCCTCAAAAATTT 830
668 GAAGTTTGTCTCCCGAGGAGATTTTGTAGGCTGTTATGATGATGTAACGCTTAGTGT 727
831 GTGAGCTCTGATCTCTTAAGGCTTATCAAGCTTGCAGTACGAGGAAGATTAAGCTT 890
728 GCGAGCGGAGGATCAAAAGGGTCTTTCAAGCTTGTCTGTCGACGGAAGATCAAGGTT 787
891 TCACTGGAATTGATGATCCATACGAAACCAATTAATGATGATAGTAAATTAAGATGA 950

ORIGIN

Query Match 24.4%; Score 296.4; DB 8; Length 1068;
Best Local Similarity 68.0%; Pred. No. 4.1e-61;
Matches 430; Conservative 0; Mismatches 196; Indels 6; Gaps 1;
411 TGTGAAACATTGGGAAATCGACTAATATTTATGGACAAATTCCTTGGATGGACAACTCG 470
308 TGTCTACGGTCGGAACCTCGACAAATATAAAGTGGCATGAATGTCTCTGTGAGAAAGTTG 367
471 ATAGACAGAAATGTCTGGACAAAAGGCTGTCTGTATGGATACAGCACTCACTGGTT 530
368 ATAGACAGAGATGTCTTGATCAGAAGGATGTGTGATTTGGGTCAACGGTCTTAGTGGTT 427
531 CAGGGAAGAGTACTCTTGCATGTGCACCTGAGTGTGATTCAGAGGCCACCTCA 590
428 CAGGGAAGAGTACTTTGGCTTGTGCTTGAATCAGATGTTGATCAAAAGGGAGCTTT 487
591 COTATGTACTTGAATGGTGACACCTCAGACATGGCTTAATAGAGATTTAAGCTTTAAGG 650
488 GTTATATTCTTGAATGGTGATATGTTAGGCATGGCTTAAACCGTATCTTAGCTTTAAAG 547
651 CAGAGACCGTCGAGAAATATACGAGAGTTGGTGAAGTGGCAAGCTTTTGTGCTATG 710
548 CTGAGGATCGTCGAGAGATATTCGTAGAGTTGGAGAGTTGCTTAAGCTTTTTCGGATG 607
711 CTGGTGTCTATGCAATGCTAGCTGATATCTCCATACAGGAGAGATCGTGTATGATGCC 770
608 CTGGAATATCTGCATTTGGAGTTTGAATCTCTTATAGACATAGGAGCGTTGTC 667
771 GTGCTCTACTTCCATCTTAATCTTATTTAAGAGTATTTATGATTTGCCCTCAAAAATTT 830
668 GAAGTTTGTCTCCCGAGGAGATTTTGTAGGCTGTTATGATGATGTAACGCTTAGTGT 727
831 GTGAGCTCTGATCTCTTAAGGCTTATCAAGCTTGCAGTACGAGGAAGATTAAGCTT 890
728 GCGAGCGGAGGATCAAAAGGGTCTTTCAAGCTTGTCTGTCGACGGAAGATCAAGGTT 787
891 TCACTGGAATTGATGATCCATACGAAACCAATTAATGATGATAGTAAATTAAGATGA 950

788 TTACCGGATCGATGACACCTTACGAGCCACCATTCGAG-----ATTCTCTAG 841
951 AAGATGAGGATGCCCTTACCCAAAGCAATGGCCCAAGATTTCTATGCTACCTTGAAG 1010
842 GACGTGAAGGAGGAATCTTCTCCTATCGAATGGCGGAAGGTCGCGATATCTAGATA 901
1011 AAAACGGATATTTCGAAGCTTAGTATATATAT 1042
902 ACAAGGTTATCTTCAAGCATACATATCTTCT 933
ATU05238 1077 bp mRNA linear PLN 23-JUN-1998
Arabidopsis thaliana APS kinase mRNA, complete cds.
U05238
VERSION U05238.1 GI:450234
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1077)
AUTHORS Jain, A. and Leustek, T.
TITLE A cDNA clone for 5'-adenylylphosphosulfate kinase from Arabidopsis
thaliana
JOURNAL Plant Physiol. 105 (2), 771-772 (1994)
MEDLINE 94345022
PUBMED 8066145
REFERENCE 2 (bases 1 to 1077)
AUTHORS Lee, S. and Leustek, T.
TITLE APS kinase from Arabidopsis thaliana: genomic organization,
expression, and kinetic analysis of the recombinant enzyme
JOURNAL Biochem. Biophys. Res. Commun. 247 (1), 171-175 (1998)
MEDLINE 98300303
PUBMED 9636674
REFERENCE 3 (bases 1 to 1077)
AUTHORS Leustek, T.
TITLE Direct Submission
JOURNAL Submitted (18-JAN-1994) Thomas Leustek, Center for Agricultural
Molec. Biology, Rutgers University, Cook College, College Farm
Road, New Brunswick, NJ 08903, USA
FEATURES
source Location/Qualifiers
1. .1077
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="ATK"
/tissue_type="whole plant"
33. .863
/EC_number="2.7.1.25"
/function="phosphorylates 5'-adenylylsulfate on
3'-hydroxyl group"
/note="similar to yeast APS kinase (MET14): SwissProt
Accession Number Q02196; similar to Escherichia coli APS
kinase (CysC): SwissProt Accession Number P23846;
adenosine 5'-phosphosulfate kinase"
/codon_start=1
/evidence="experimental
product="APS kinase"
/protein_id="AAC50035.1"
/db_xref="GI:450235"
transit_peptide 33. .143
/note="potential chloroplast transit peptide with a
protease cleavage site"

2

```
QY 531 CAGGGAAGAGTACTCTTGCATGTGCACTGAGTCGTGAGTTGCAATGCGAGAGCCACCTCA 590
Db 427 CAGGGAAGAGTACTCTTGGCTTGTGCTTTGAATCAGATGTTGATCAAAAGGGGAGCTTT 486
QY 591 CCGTATCTTCTGATGTCACACCTCAGACATGGCTAAATAGAGATTAGCTTTAAGG 650
Db 487 GTTATATCTTGTGATGATAATGTTAGGCATGGCTTAAACCGTGATCTTAGCTTTAAG 546
QY 651 CAGAAGACCGTCGAGAAAATATACGAAGAGTTGGTGAAGTGCAGAAAGCTTTTGTGATG 710
Db 547 CTGAGGATCGTCAGAGATATTCGTAGAGTTGGAGAGGTTGCTAAGCTTTTTCGGATG 606
QY 711 CTGGTGTCAATGATGCTAGCTGATATCTTCATATCCATACAGAGAGATCGTGATGATGCC 770
Db 607 CTGGAATAATCTGCATTCGAGTTGCGAGTTGATATCTCTTATAGAACAGATAGGACGCTTGC 666
QY 771 GTGCTCTACTCCACATCTCACTTTTATTTAGATATTTATTGATTTGCCCTAAAAATTT 830
Db 667 GAAGTTTGTCTCCCGAGGAGATTTTGTGTAGGTGTTGATGATGTACGCTTAGTGT 726
QY 831 GTGAAGCTCGTGATCTTAAAGGCTTATACAAAGCTTTCACGCTGACAGAAAGATTAAAGTT 890
Db 727 GCGAGCGAGGATCCAAAGGCTTTTACAGCTTGTCTGTCGACAGAAAGATCAAGGTT 786
QY 891 TCATCGGAATTGATGATCATACGAACCAATTAATGTTGATAGATGTAATTAAGATGA 950
Db 787 TTACCGGATCGATGACCTTACGAGCCACCATTTGAACTGCAG-----ATTCTCTAG 840
QY 951 AGATGAGGAATGCCCTTCCACCAAGCATGCGCAAGCAAGTTCTATGCTACCTTGAAG 1010
Db 841 GACGTGAAGGAGGAATCTCTCTATCGAAATGGCGGAAAGTCTGCGATCTTAGATA 900
QY 1011 AAAACGGATATTGCAAGCTTAGTATATGAT 1042
Db 901 ACAAGGTTATCTTCAAGCATTAACATCTTCT 932
```

RESULT 8

```
ATCLAPSK
LOCUS ATCLAPSK 1185 bp mRNA linear PLN 05-FEB-1998
DEFINITION A.thaliana (L.Heynh.) chloroplast mRNA for recombinant APS-kinase.
```

```
ACCESSION X75782
```

```
VERSION X75782.1 GI:414736
```

```
KEYWORDS APS-kinase; ATP:adenylylsulfate-3'-phosphotransferase.
```

```
SOURCE Arabidopsis thaliana (thale cress)
```

```
ORGANISM Arabidopsis thaliana
```

```
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
```

```
REFERENCE 1
```

```
AUTHORS Arz,H.E., Gisselmann,G., Schiffmann,S. and Schwenn,J.D.
```

```
TITLE A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana
```

```
JOURNAL Biochim. Biophys. Acta 1218 (3), 447-452 (1994)
```

```
MEDLINE 94325358
```

```
PUBMED 8049272
```

```
REMARK (sites)
```

```
REFERENCE 2 (bases 1 to 1185)
```

```
AUTHORS Arz,H.E., Gisselmann,G., Schiffmann,S. and Schwenn,J.D.
```

```
TITLE A chloroplast APS-kinase cDNA from Arabidopsis thaliana
```

```
JOURNAL Unpublished
```

```
REFERENCE 3 (bases 1 to 1185)
```

```
AUTHORS Schwenn,J.D.
```

```
TITLE Direct Submission
```

```
JOURNAL Submitted (03-NOV-1993) Schwenn J. D., Ruhr University Bochum,
```

```
Biolog, Universitaetsstr. 150, 44780 Bochum, Germany
```

```
Location/Qualifiers
```

```
1..1185
```

```
/organism="Arabidopsis thaliana"
```

```
/mol_type="mRNA"
```

```
/db_xref="taxon:3702"
```

```
/clone="FM239x14cDNA"
```

```
/tissue_type="leaf"
```

```
/clone_lib="cDNA in lambda ZapII"
```

FEATURES

source

```
CDS
/dev_stage="vegetative"
42..872
/EC_number="2.7.1.25"
/function="phosphorylates 3'-OH group of adenylylsulfate"
/standard_name="ATP:adenylylsulfate-3'-phosphotransferase"
/citation=[2]
/codon_start=1
/product="APS kinase"
/protein_id="CAA53426.1"
/db_xref="GI:414737"
/db_xref="COA_Q43295"
/db_xref="UniProt/Swiss-Prot:Q43295"
/translation="MIAAGAKSLGLSWASPKGIFDMSMSNSRSVVVRAVCVMDGS
QTLNHNKSGSPEVKSNGHTGQKGPLSTVGNSTNWKHSCSVKVDQRLLDQKGC
VIWYGLSGSGKSLACALNOMLYQKGLCYILDNDVNRHGLNDRDLSPKABRANIR
RVGVAKLFPADAGIICIASLISPYRTDRDACSLLPEGDFVEVMDVPLSVCEARDPK
GLYKLARAGKIKGFTGIDDDPYEPPLNCEISLGRSGGTSPIEMAEKVGWYLDNKGVLQA
"
162..869
/mat_peptide
/product="APS-kinase"
/EC_number="2.7.1.25"
/notes="putative mature peptide only - gene product
starting with ATG at position 162 is enzymatically active,
but maturation in the chloroplast has not yet been shown.
Homology to sequences from E. coli & ye"
/citation=[1]
/citation=[2]
/transit_peptide 42..153
/notes="putative transit peptide with a processing site
suitable for stromal protease; ref: Gavel & Heijne FEBS
Lett. 261, 455, 1990"
984..1012
stem_loop
polyA_site
985..990
stem_loop
1092..1120
ORIGIN
```

```
Query Match 24.4%; Score 296.4; DB 8; Length 1185;
Best Local Similarity 68.0%; Pred. No. 4.2e-61;
Matches 430; Conservative 0; Mismatches 196; Indels 6; Gaps 1;
QY 411 TGTGAACATTTGGAAATCGACTAATATTTATGGCACAANTGCTTGATGGACAACTCG 470
Db 256 TGTCTACGGTCGGAACCTCGACAAATATAAAGTGGCATGAATGTTCTCTTGAGAAAGTTG 315
QY 471 ATAGACAGAAATTCCTGGGACAAAAGCTCTGCTGATGATAACAGACTCAGTGTGTT 530
Db 316 ATAGACAGAGATTCCTGATCGAAAGATGTGATTTGGGTCAACCGTCTTAGTGTGTT 375
QY 531 CAGGGAAGATGACTCTTTGCATGTCACGTGCTGAGTTGCTCATTTGCAGAGGCCACCTCA 590
Db 376 CAGGGAAGATGACTCTTTGGCTTGTGCTTTGAATCAGATGTTCTATCAAAAGGGAGCTTT 435
QY 591 CGTATGTAATGATGTTGATGCAACCTCAGACATGGCTTAATAGAGATTTAAGCTTTAAGG 650
Db 436 GTTATATTTCTTGTATGTTGATAATGTTAGGCATGGCTTAAACCGTATCTTAGCTTTAAG 495
QY 651 CAGAAGACCGTCGAGAAAATATACGAAGATTTGGTGAAGTGCAGAAAGCTTTTGTCTGATG 710
Db 496 CTGAGATCGTGCAGAGATTAATTCGTAGATTTGAGAGATTTGCTAAGCTTTTTCGGATG 555
QY 711 CTGGTGTCAATGCAATGCTAGCTTTGATATCTCCATACAGAGAGATCGTGATGATGCC 770
Db 556 CTGGAAATACTGCATTTGCGAGTTTGATATCTCTTATAGAACAGATAGGACGCTTCTC 615
QY 771 GTGCTCTACTTCCACATCTTAACCTTTATTGAAGTATTTATTGATTTGCCCTAAAAATTT 830
Db 616 GAAGTTTGTCTCCCGAGGGAGATTTTGTGTAGGTGTTTCATGGATGTACCGCTTAGTGT 675
QY 831 GTGAAGCTCGTGATCTCTAAAGCGCTTATCAAGCTTTGCACGTACAGGAAGAGATTAAAGTT 890
Db 676 GCGAGGCGAGGGATCCAAAGGGTCTTTTACAAGCTTTGCTCGTGCAGAGAAAGATCAAAGTT 735
QY 891 TCACTGGAAATTGATGATCCATACGAACCAATTAATTTGTTGATAGTAAATTAAGATGA 950
```

```
Db      736 TTACCGGATCGATGACCCCTTACGAGCCACCAATTGAATGCGAG-----ATTTCTCTAG 789
Qy      951 AAGATGAGGATGCCCTTACCCAAAGCAATGGCCCAAGCAAGTTCTATGCTACCTTGAAG 1010
Db      790 GACGTGAAGGAGAACTTCTCTATCGAAATGGCGGAAAGGTCGTCGGATACTTAGATA 849
Qy      1011 AAAACGGATATTTCGAAGCTTAGTATGTAT 1042
Db      850 ACAAGGGTTATCTTCAAGCAATACTACTTCT 881

RESULT 9
LOCUS      CQ805858                831 bp      DNA      linear      PAT 10-MAY-2004
DEFINITION Sequence 2269 from Patent WO2004035798.
ACCESSION  CQ805858
VERSION    CQ805858.1   GI:47111563
KEYWORDS
SOURCE     Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE  1
AUTHORS   Inze,D.; de Veylder,L. and Vlieghe,K.
TITLE      Identification of novel e2f target genes and use thereof
JOURNAL    Patent: WO 2004035798-A 2269 29-APR-2004;
CropDesign N.V. (BE)
FEATURES   source
            1..831
            /organism="Arabidopsis thaliana"
            /mol_type="unassigned DNA"
            /db_xref="taxon:3702"

ORIGIN
Query Match      24.2%; Score 294.4; DB 6; Length 831;
Best Local Similarity 68.3%; Pred. No. 1.2e-60;
Matches 425; Conservative 0; Mismatches 191; Indels 6; Gaps 1;

Qy      411 TGTCAACATTTGGGAAATCGACTAATATTTATGGCACAATTCGTTGATGGACAATCTG 470
Db      215 TGTCTACCGTCGGAAACTCGACAAATATAAAGTGGCATGAATGTTCTCTTGAGAAAGTTG 274
Qy      471 ATAGACAGAAATTCGTCGGACAAAAGGCTGTGCTATGGATAACAGCACTCACTGTGTT 530
Db      275 ATAGACAGAGATTGCTTGATCAGAAGAGATGTGTGATTTGGTCAACCGTCTTAGTGGTT 334
Qy      531 CAGGAAAAGTACTCTTCGCATGTGCATCGAGTCGTGAGTTGCATTCGACAGGCCACCTCA 590
Db      335 CAGGGAAGAGTACTTTGGCTTGTGCTTTGAATCAGATGTTGTATCAAAAGGGGAAGCTTT 394
Qy      591 CGTATGTACTTGTATGGTCACACCTCAGACATGGCTTAATAGAGATTTAAGCTTTAAGG 650
Db      395 GTTATATCTTGTATGGTGATAATCTGTAGAGTTGGAGAGTTGCTAAGCTTTTGGCGATG 454
Qy      651 CAGAAGACCGTCGACAAAATATACGAAGATTGGTGAAGTGGCAAGCTTTTGTCTGATG 710
Db      455 CTGAGATCGTGCAGAGATATCTGTAGAGTTGGAGAGTTGCTAAGCTTTTGGCGATG 514
Qy      711 CTGGTGTCTATATGCAATTCAGTTGATATCTCCATACAGGAGATCGTGTATGATGCC 770
Db      515 CTGGAATAATCTGCATTCGAGTTTGAATCTCTCTTATAGAACAGATAGGACCGTTGTC 574
Qy      771 GTGCTCTACTTCCACATTTCTAATCTTATGAAGTATTTATTTGATTTGCCCTAAAAATTT 830
Db      575 GAAGTTTCTCCCGAGGGAGATTTTGTGAGGTGTTTCATGGATGTACCGCTTAGTGTTT 634
Qy      831 GTGAAGCTCGTGATCCTTAAAGCCCTATACAAAGCTTGCACGTACAGGAAGATTTAAAGTT 890
Db      635 GCGAGCGGAGGATCCAAAGGCTCTTACAGAGCTTGCTCGTGCAGGAAGATCAAAGTT 694
Qy      891 TCATCGGAATTGATGATCCATACGAACCAACCAATTAATGGTGGAGATAGTAATTAAGATGA 950
```

```
Db      695 TTACCGGATCGATGACCCCTTACGAGCCACCAATTGAATGCGAG-----ATTTCTCTAG 748
Qy      951 AAGATGAGGATGCCCTTACCCAAAGCAATGGCCCAAGCAAGTTCTATGCTACCTTGAAG 1010
Db      749 GACGTGAAGGAGAACTTCTCTATCGAAATGGCGGAAAGGTCGTCGGATACTTAGATA 808
Qy      1011 AAAACGGATATTTCGAAGCTTA 1032
Db      809 ACAAGGGTTATCTTCAAGCATA 830

RESULT 10
LOCUS      AX412633                831 bp      DNA      linear      PAT 14-JUN-2002
DEFINITION Sequence 397 from Patent WO0222675.
ACCESSION  AX412633
VERSION    AX412633.1   GI:21445091
KEYWORDS
SOURCE     Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE  1
AUTHORS   Glazebrook,J., Wang,X., Dangl,J.L., Eulgem,T. and Zhu,T.
TITLE      Plant genes, the expression of which are altered by pathogen
            infection
JOURNAL    Patent: WO 0222675-A 397 21-MAR-2002;
Syngenta Participations AG (CH) ; UNIVERSITY OF NORTH CAROLINA AT
CHAPEL HILL (US) ; Glazebrook, Jan (US) ; Wang, Xun (US) ; Dangl,
Jeffrey L. (US) ; Eulgem, Thomas (US)
FEATURES   Location/Qualifiers
            source
            1..831
            /organism="Arabidopsis thaliana"
            /mol_type="unassigned DNA"
            /db_xref="taxon:3702"

ORIGIN
Query Match      24.2%; Score 294.4; DB 6; Length 831;
Best Local Similarity 68.3%; Pred. No. 1.2e-60;
Matches 425; Conservative 0; Mismatches 191; Indels 6; Gaps 1;

Qy      411 TGTCCGAACATTTGGGAAATCGACTAATATTTATGGCACAATTCGTTGATGGACAATCTG 470
Db      215 TGTCTACCGTCGGAAACTCGACAAATATAAAGTGGCATGAATGTTCTCTTGAGAAAGTTG 274
Qy      471 ATAGACAGAAATTCGTCGGACAAAAGGCTGTGCTATGGATAACAGCACTCACTGTGTT 530
Db      275 ATAGACAGAGATTGCTTGATCAGAAGAGATGTGTGATTTGGGTCAACCGTCTTAGTGGTT 334
Qy      531 CAGGAAAAGTACTCTTTCGCATGTGCATCGAGTCGTGAGTTGCATTCGACAGGCCACCTCA 590
Db      335 CAGGGAAGAGTACTTTGGCTTGTGCTTTGAATCAGATGTTGTATCAAAAGGGGAAGCTTT 394
Qy      591 CGTATGTACTTGTATGGTGAACAACCTCAGACATGGCTTAAATAGAGATTTAAGCTTTAAGG 650
Db      395 GTTATATCTTGTATGGTGATAATCTGTAGCATGGCTTAAACCGGTGATCTTAGCTTTAAAG 454
Qy      651 CAGAAGACCGTCGACAAAATATACGAAGATTGGTGAAGTGGCAAGCTTTTGTCTGATG 710
Db      455 CTGAGGATCGTGCAGAGATATTCGTAGAGTTGGAGAGTTGCTAAGCTTTTGGCGATG 514
Qy      711 CTGGTGTCTATATGCAATTCAGTTGATATCTCCATACAGGAGATCGTGTATGATGCC 770
Db      515 CTGGAATAATCTGCATTCGAGTTTGAATCTCTCTTATAGAACAGATAGGACCGTTGTC 574
Qy      771 GTGCTCTACTTCCACATTTCTAATCTTATGAAGTATTTATTTGATTTGCCCTAAAAATTT 830
Db      575 GAAGTTTCTCCCGAGGGAGATTTTGTGAGGTGTTTCATGGATGTACCGCTTAGTGTTT 634
Qy      831 GTGAAGCTCGTGATCCTTAAAGCCCTATACAAAGCTTGCACGTACAGGAAGATTTAAAGTT 890
```


Db 635 GCGAGCGGAGGATCCAAAGGGTCTTTTACAAGCTTGCTCGTGCGAGAAAGATCAAAGGTT 694

Qy 891 TCACTGGAATTGATGATCCATACGAACCAACCAATTAATGCTGAGATAGTAATTAAGATGA 950

Db 695 TTACCGGGATCGATGACCTTTACGAGCCACCATTTGAATGCGAG-----ATTCTCTAG 748

Qy 951 AGATGAGGAATGCCCTTCCACCAAGCAATGCGCAAGCAAGTTCTATGCTACCTTTGAAG 1010

Db 749 GACGTGAAGGAGGAACCTTCTCTATCGAAATGCGGAAAGGTCGTCGGATACTTAGATA 808

Qy 1011 AAAACGGATATTGCAAGCTTA 1032

Db 809 ACAAGGGTTATCTTCAAGCATA 830

RESULT 11

AX506015

LOCUS AX506015 831 bp DNA linear PAT 27-SEP-2002

DEFINITION Sequence 710 from Patent WO0216655.

ACCESSION AX506015

VERSION AX506015.1 GI:23387252

KEYWORDS Arabidopsis thaliana (thale cress)

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 Harper, J.F., Kreps, J., Wang, X. and Zhu, T.

AUTHORS Stress-regulated genes of plants, transgenic plants containing same, and methods of use

TITLE Patent: WO 0216655-A 710 28-FEB-2002;

JOURNAL The Scripps Research Institute (US); Syngenta Participations AG (CH)

FEATURES

source Location/Qualifiers

1..831

/organism="Arabidopsis thaliana"

/mol_type="unassigned DNA"

/db_xref="taxon:3702"

ORIGIN

Query Match 24.2%; Score 294.4; DB 6; Length 831;

Best Local Similarity 68.3%; Pred. No. 1.2e-60;

Matches 425; Conservative 0; Mismatches 191; Indels 6; Gaps 1;

Qy 411 TCTCGAACATTTGGGAATCGACTAATATTTATTTATGCGCAATTCCTTGATGGACATCTG 470

Db 215 TGTCTACGTCGGAACCTCGACAAATATAAAGTGGCATGAATGTTCTGTGAGAAAGTTG 274

Qy 471 ATAGACAGAAATTTGCTGGACAAAAGGCTGTGCTGATGGATAACAGGACTCAGTGGTT 530

Db 275 ATAGACAGAGATGCTTGTATCAGAAAGATGTGTGTTGGTCCACGGTCTTAGTGGTT 334

Qy 531 CAGGGAAGATGACTCTTTCATGTGCACTGAGTTCGATTCGATTCGAGAGCCACCTCA 590

Db 335 CAGGGAAGATGACTTTGGCTTGTGCTTTCGAATCAGATGTTGTATCAAAGGGGAGCTTT 394

Qy 591 CGTATGTACTTGTATGTTGACACCTCAGACATGGCCCTAATAGAGATTTAAGCTTTAAG 650

Db 395 GTTATATTCTGATGTTGATATGTTAGGCATGGCTTAAACCGTGTATCTTAGCTTTAAG 454

Qy 651 CAGAGACCGTGCAGAAAATATACGAAGATTTGGTGAAGTGCAGAAAGCTTTTTCGTGATG 710

Db 455 CTGAGGATCGTCAGAGAAATATTCGTAAGATTTGGAGAGTTCCTAAGCTTTTTCGGATG 514

Qy 711 CTGGTGTATATGCAATTTGCTAGCTTGCATATCTCCATACAGAGAGATCGTGTATGCC 770

Db 515 CTGGAATAATCTGCATTGCGAGTTGATATCTCTTATAGACAGATAGGAGCGCTTGTG 574

Qy 771 GTGCTCTACTTCGACATTTCTAATTTTATTTGAAGTATTTATGATTTGCCCTTAAATTT 830

Db 575 GAAGTTTGTCCCGGAGGAGATTTTGTGAGGTGTTTCAATGATGTACCGCTTAGTGTGTT 634

Qy 831 GTGAAGCTCGTGATCCTAAAGCCCTATACAAGCTTGCACGTACAGGAAGATTAAGGTT 890

Db 635 GCGAGCGGAGGATCCAAAGGGTCTTTTACAAGCTTGCCTCGTCAGGAAGATCAAAGGTT 694

Qy 891 TCATCGGAATGATGATCCATACGAACCAACCAATTAATGCTGAGATAGTAATTAAGATGA 950

Db 695 TTACCGGGATCGATGACCTTTACGAGCCACCATTTGAATGCGAG-----ATTCTCTAG 748

Qy 951 AAGATGAGGAATGCCCTTCCACCAAGCAATGCGCAAGCAAGTTCTATGCTACCTTTGAAG 1010

Db 749 GACGTGAAGGAGGAACCTTCTCTATCGAAATGCGGAAAGGTCGTCGGATACTTAGATA 808

Qy 1011 AAAACGGATATTGCAAGCTTA 1032

Db 809 ACAAGGGTTATCTTCAAGCATA 830

RESULT 12

AX132010

LOCUS Arabidopsis thaliana At2g14750/F26C24.11 mRNA, complete cds.

DEFINITION Arabidopsis thaliana (thale cress)

ACCESSION AY132010

VERSION AY132010.1 GI:22135772

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 831)

AUTHORS Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A., and Ecker, J.R.

TITLE Arabidopsis ORF clones

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 831)

AUTHORS Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A., and Ecker, J.R.

TITLE Direct Submission

JOURNAL Submitted (26-JUN-2002) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGECC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Banh, J., Bowser, L., Chang, E., Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

Location/Qualifiers

source	1. .831	LOCUS	BT005193	964 bp	linear	PLN 04-MAR-2003
	/organism="Arabidopsis thaliana"	DEFINITION	Arabidopsis thaliana clone U20575 putative adenylylsulfate kinase (At5g67520) mRNA, complete cds.			
	/mol_type="mRNA"	ACCESSION	BT005193			
	/db_xref="taxon:3702"	VERSION	BT005193.1	GI:28827763		
	/chromosome="2"	KEYWORDS	FLJ CDNA.			
	/clone="U13881"	SOURCE	Arabidopsis thaliana (thale cress)			
	/ecotype="Columbia"	ORGANISM	Arabidopsis thaliana			
	/note="This clone is in pUNI 51"					
CDS	1. .831	REFERENCE	1 (bases 1 to 964)			
	/note="putative adenosine phosphosulfate kinase"	AUTHORS	Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.			
	/codon_start=1	TITLE	Arabidopsis Open Reading Frame (ORF) Clones			
	/product="At2g14750/F26C24.11"	JOURNAL	Unpublished			
	/protein_id="AA091043.1"	REFERENCE	2 (bases 1 to 964)			
	/db_xref="GI:22135773"	AUTHORS	Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.			
	/translational="MIAAGKLLGLSWASPKGIPDSMSMNSRSVVVVRACVSMOGS QTLSHNKSGISPEVKSINHGQKQPLSTVGNSTNIRWHECSVEKVDRLQLPQKGC VTWLGLSGSGSTLACALQMLYQKGLCYLTDGDNVHGLNDRDLSFKADRAENIR RVGEVAKLPADAGIICIASLIPYTRDRAKSLLEPGDFVEVFMVPLSVCEARDPK GLYKLARACKIKGFTGIDDPYEPPLNCEISLGRGGTSPIEWAEKVVGYLNDKNGVLOA "	TITLE	Arabidopsis Full-length cDNA: Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J., Hayashizaki,Y. and Shinozaki,K.			
	/note="putative adenosine phosphosulfate kinase"	JOURNAL	Submitted (04-MAR-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA			
	/codon_start=1	COMMENT	The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : RIKEN Arabidopsis Full-length cDNA): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J., Hayashizaki,Y. and Shinozaki,K.			
	/product="At2g14750/F26C24.11"		The Salk, Stanford, PGECC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAPL cDNAs: Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.			
	/translational="MDVAAMARCVGCYSPAFGESEHRLSERRFLKLSSTNSDPA GSKLRLGRIHMRYSFRPQWAKDESISSRGCTQKINGKQNIWHDCPVTKSDRQ ELIKQKQCVITWITGLSGSGKSSLACALSRLNHRGLSYILDGDNVHGLNDSLSEFA "		Yamada,K. (SSP/PGECC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP /PGECC) contributed equally to this work as PIs.			
	/codon_start=1	FEATURES	Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.			
	/product="putative adenylylsulfate kinase"	source	Location/Qualifiers			
	/protein_id="AA050726.1"		1. .964			
	/db_xref="GI:28827764"		/organism="Arabidopsis thaliana"			
			/mol_type="mRNA"			
			/db_xref="taxon:3702"			
			/chromosome="5"			
			/clone="U20575"			
			/ecotype="Columbia"			
			/note="This clone is in pUNI 51."			
gene	1. .964		/gene="At5g67520"			
CDS	1. .933		/gene="At5g67520"			
			/codon_start=1			
			/evidence=experimental			
			/product="putative adenylylsulfate kinase"			
			/protein_id="AA050726.1"			
			/db_xref="GI:28827764"			
			/translation="MDVAAMARCVGCYSPAFGESEHRLSERRFLKLSSTNSDPA GSKLRLGRIHMRYSFRPQWAKDESISSRGCTQKINGKQNIWHDCPVTKSDRQ ELIKQKQCVITWITGLSGSGKSSLACALSRLNHRGLSYILDGDNVHGLNDSLSEFA "			
RESULT 13						
BT005193						

DDRAENIRRVGEVAKLFADSGIICIASLISPYRIERACRALLPQDGFIEFMDVPLH
VCEARDPKGLYRABAGKIKGFTGVDDPYEAPLDCEIVIQNSRDKGLSSSSSSSPS
SSSSSLCEMADIVSVYLDQNGYKXKHSKSRNCM

misc_difference 768

/gene="At5g67520"

/note="compared to genomic sequence"

/replaces="a"

934..964

/gene="At5g67520"

3'UTR

ORIGIN

Query Match 23.7%; Score 288.4; DB 8; Length 964;

Best Local Similarity 70.0%; Pred. No. 3.6e-59;

Matches 388; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

QY 416 AACATTGGCAATCGACTAATATTTATGGCAATGCTTGCATTCGACAACTCATAGA 475
Db 244 ATCAATGGAAAGCAAAAGAACATGTCGGCATGATGTCGGTTACTTAAATCCGACAGG 303
QY 476 CAGAAATTCCTGGGCAAAAAGGCTGTGCTATGATAACAGGACTCAGTGGTTTCAGGG 535
Db 304 CAAGAATTAATTAAGCAAAAGGATGTGATTTGATTTACTGGCTTAAGTGTTCAGGT 363
QY 536 AAAAGTACTCTTGATGTCGACTGATGCTGATGATGCTGATGATGCTGATGATGCTGAT 595
Db 364 AAAAGTAGTCTGGCATGTCTCTTAGTCGAGCTTTGCACAATCGTGGAAAGCTTTCGTAT 423
QY 596 GTACTTGATGTCGACAACTCAGACATGCGCTTAATAGATGATTTAAGCTTTAAGCAGNA 655
Db 424 ATACTTGATGGTGACAAATGTCGACATGCTTTAAACAGCGATCTTAGTTTGAAGCAGAT 483
QY 656 GACCGTGCAAAAATATACAGAGATGTGTGAAGTGGCAAGCTTTTGTCTGATGCTGT 715
Db 484 GATCAGCTGAACAATTCGAAGATGTGTGAAGTGGCTTAAGTGTTCAGATTTCTGT 543
QY 716 GTCATATGATGCTAGCTTGATATCTCCATACAGAGAGATCGTATGATGATGCTGCT 775
Db 544 ATATCTGTATTGCAAGTTAATATCTCTTACCGATAGACGAGCTGCTGCGCTGCA 603
QY 776 CTACTTCCATCTTAACCTTTATGAGTATTTATGATTTGCTCCCTTAAATTTGTGA 835
Db 604 TTATTACCAAGAGGATTTCAATGAGGTATTTATGATGCTGCTCACTCCATGTTGTGA 663
QY 836 GCTGTGATCTCTAAAGGCTTATACAGCTTGACAGGAAAGATTAAGGTTTCACT 895
Db 664 GCTAGAGATCCAAAGGCTTATACAACTGACGCGCTGGTAAATCAAGGTTTACA 723
QY 896 GGAATGATGATCCATACGACCAACCAATTAATGTTGATAGTAAATGAAGATGAAGAT 955
Db 724 CGAGTAGATGATCCATATGAAGGCTTTGGATTCGAGATTTGTATACAAACAGTAGA 783
QY 956 GAGGAATGCCCTTC 969
Db 784 GACAAGGGGCTTTC 797

RESULT 14

AY085031

LOCUS

Arabidopsis thaliana clone 1133 bp mRNA linear PLN 14-APR-2003

ACCESSION AY085031

VERSION AY085031.1 GI:21403741

KEYWORDS FLI CDNA

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1133)

Haas,B.J., Volkovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,

Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.B.

Full-length messenger RNA sequences greatly improve genome

annotation

JOURNAL

MEEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Source

CDS

ORIGIN

Query Match

Best Local

Similarity

Matches

Conservative

Score

DB

Length

Mismatches

Indels

Gaps

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

Genome Biol. 3 (6), RESEARCH0029 (2002)

22088475

12093376

2 (bases 1 to 1133)

Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and

Feldmann,K.

Full-length cDNA from Arabidopsis thaliana

Unpublished

3 (bases 1 to 1133)

Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and

Feldmann,K.

Direct Submission

Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,

Malibu, CA 90265, USA

This clone sequence is one of 5,000 Ceres full-length cDNAs made

available to TIGR and Genbank. The following quality assessment of

this set was done by comparison with known proteins: two percent of

the clones are estimated to be 5'-truncated; less than one percent

are 3'-truncated; approximately two percent represent alternative

splice variants, including unspliced introns and spliced exons; one

percent may contain premature stop codons; five percent may have

frame shifts in a coding region. A sequence is considered to be

5'-truncated if it lacks the translation initiation start (ATG). A

sequence is considered to be 3'-truncated if it lacks the

C-terminal end of the encoded protein. Please note that these cDNA

sequences are derived from the Ws or Laer ecotypes and therefore

may contain polymorphisms when compared to sequences from Col-0.

Genet carried out the library production and sequencing of the

full-length clones. Ceres, Inc. carried out the clustering of the

5' sequences, selection of clones, and sequence assembly.

Location/Qualifiers

1..1133

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/clone="125039"

111..1028

/codon_start=1

/product="adenylsulfate kinase-like protein"

/protein_id="AA061589.1"

/db_xref="GI:21537248"

/translation="MARCVGRVYSPAFGESEHRLSRFLKSSSTNSDPAGSKSL

KLRGKIRMSYFRPIKADSESISSRGETQKINGKNIVHIDCPVTKSRQELIKH

KGCVITWTGLSGSKSLACALSRALHNRKLSVILGDGNVHGLNSDLSPEADRAE

NIRVGEVAKLFADSGIICIASLISPYRIERACRALLPQDGFIEFMDVPLHVCAR

DPKGLYRABAGKIKGFTGVDDPYEAPLDCEIVIQNSRDKGLSSSSSSSSSSSSSS

LCENADIVSVYLDQNGYKXKHSKSRDCM"

Qy	896	GGAAATTGATGATCCATACGAACCAACCAATTAATGTGAGATAGTAATTAAGATGAAAGAT	955
Db	824	GGAGTAGATGATCCATATGAAGGCCCTTTGGATTGGAGATTGTGTACAAACACAGTAGA	883
Qy	956	GAGGAATGCCCTTC	969
Db	884	GACAGGGGCTTC	897

Search completed: September 9, 2005, 11:35:54
Job time : 5333 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 9, 2005, 08:53:59 ; Search time 681 Seconds
(without alignments)
10579.038 Million cell updates/sec

Title: US-10-829-432-3

Perfect score: 1217

Sequence: 1 ggcgcgttcatttcacaa.....aaaaaaaaaaaaaaaaaaaaa 1217

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1217	100.0	1217	3	Aaz50160 Corn Aden
2	359.6	29.5	928	3	Aaz50163 Wheat Ade
3	350.2	28.8	378	12	Adj44283 Plant cDN
4	338.8	27.8	936	3	Aaz50162 Soybean A
5	320.4	26.3	627	3	Aac44691 Arabidops
6	316.2	26.0	890	3	Aaz50159 Corn Aden
7	296.4	24.4	1068	3	Aac35266 Arabidops
8	296.4	24.4	1175	3	Aac36588 Arabidops
9	296.4	24.4	1183	3	Aac45659 Arabidops
10	294.4	24.2	831	6	Abz12905 Arabidops
11	294.4	24.2	831	6	Adg87955 A. thalia
12	294.4	24.2	831	12	Adn74374 Thale cre
13	290	23.8	1130	3	Aac49070 Arabidops
14	288.4	23.7	1133	3	Aac36627 Arabidops
15	264.8	21.8	687	12	Adj39869 Plant cDN
16	263.6	21.7	917	3	Aaz98337 A. thalia
17	263.6	21.7	948	3	Aac42294 Arabidops
18	245	20.1	262	7	Adg66664 Corn seed
19	233.4	19.2	595	12	Adj43843 Plant cDN
20	225	18.5	714	3	Aac51190 Arabidops

21	198.2	16.3	483	3	AAC51315	Aac51315 Arabidops
22	197.2	16.2	431	3	Aaz50161	Aaz50161 Rice Aden
23	188.6	15.5	328	12	ADP92312	Adp92312 Cotton ex
24	179	14.7	597	8	ACA47863	ACA47863 Prokaryot
25	170.2	14.0	603	8	ACA27590	ACA27590 Prokaryot
26	168.6	13.9	642	8	ACA53672	ACA53672 Prokaryot
27	162.8	13.4	521	3	Aaz50164	Aaz50164 Wheat Ade
28	161.4	13.3	600	13	ADT42933	Adt42933 Bacterial
29	159.6	13.1	615	6	ABN92217	Abn92217 Staphyloc
30	158.6	13.1	615	13	ADS02000	Adg02000 Staphyloc
31	158.2	13.0	597	8	ACA46934	ACA46934 Prokaryot
32	156.4	12.9	609	13	ADS3268	Adg33268 Non-ScSSU
33	156.4	12.9	1160	3	AAA08597	Aaa08597 Saccharom
34	153.2	12.6	2014	10	ADD43524	Add43524 Human cDN
35	153.2	12.6	2014	12	ADK67851	Adk67851 Phosphoad
36	153.2	12.6	2014	12	ADK67854	Adk67854 Phosphoad
37	152.6	12.5	594	13	ADT48369	Adt48369 Bacterial
38	152.2	12.5	2424	12	ADK67852	Adk67852 Phosphoad
39	152	12.5	1845	10	ADD43532	Add43532 Humanopen
40	152	12.5	3774	8	ABX10363	Abx10363 DNA encod
41	152	12.5	3774	12	ADK67853	Adk67853 Phosphoad
42	152	12.5	3774	13	ADP23471	Adp23471 PRO Polyp
43	151.4	12.4	588	13	ADT41933	Adt41933 Bacterial
44	150.6	12.4	2617	3	Aaz94206	Aaz94206 Human tra
45	150.6	12.4	2617	8	ABX77649	Abx77649 Different

ALIGNMENTS

RESULT 1

ID AAZ50160 standard; cDNA; 1217 BP.

AC AAZ50160;

DT 04-MAY-2000 (first entry)

DB Corn Adenylylsulphate kinase-2 cDNA clone.

KW Adenylylsulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
3'-Phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
KW corn; clone p0016.ctscj40rb; transgenic plant; screen; antibody; ss.

OS Zea mays.

FH Key Location/Qualifiers
CDS 2..1033

FT /*tag= a

FT /product= "Corn APS kinase-2"

FT /note= "Derived from clone p0016.ctscj40rb"

XX WO200004165-A1.

XX 27-JAN-2000.

XX 13-JUL-1999; 99WO-US015809.

XX 14-JUL-1998; 99US-0092833P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Falco SC, Allen SM, Anderson SL;

DR WPI; 2000-182430/16.

XX P-PSDB; AAY44789.

PT New nucleic acid molecule and chimeric gene encoding an adenosine-5'-
phosphosulphate kinase, useful for altering expression of sulfate
PT assimilation protein in plants.

PS Claim 3; Page 30; 42pp; English.

XX

CC The present sequence is a cDNA encoding corn adenylsulphate kinase (APS
CC kinase), also known as adenosine-5' phosphosulphate kinase. This is
CC obtained from clone p0016.ctscj40rb, derived from corn pooled tassel
CC shoots, p0016 cDNA library. APS kinase is a sulphate assimilation
CC protein, that catalyses the conversion of adenosine-5' phosphosulphate
CC (APS) to 3'-phospho-adenosine-5' phosphosulphate (PAPS). This sequence is
CC used as probe and primers to identify, obtain and synthesise sulphate
CC assimilation proteins from other plants. It is also used to produce
CC transgenic plants, that are useful for altering the expression levels of
CC a sulphate assimilation protein. The APS kinase peptides are useful for
CC producing antibodies, that are used to screen and isolate cDNA clones
XX

SQ Sequence 1217 BP; 344 A; 292 C; 303 G; 278 T; 0 U; 0 Other;

Query Match 100.0%; Score 1217; DB 3; Length 1217;
Best Local Similarity 100.0%; Pred. No. 1.2e-297;
Matches 1217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTCGGTTTCATTTTCATCAATCAACACAGAACCTCTGGTCAACACAGCGCAACACCACC 60
DB 1 GCGTCGGTTTCATTTTCATCAATCAACACAGAACCTCTGGTCAACACAGCGCAACACCACC 60

QY 61 GAGCCAGCGCCCGGCCAGCCAGCGGCCCAACGCGAAGGCAACACCTCTCAGGCC 120
DB 61 GAGCCAGCGCCCGGCCAGCCAGCGGCCCAACGCGAAGGCAACACCTCTCAGGCC 120

QY 121 GAGCGGAGCGCTCGCGTCATCTCGTAAATCCACAGCGCGCTCCCGTCTCCACGG 180
DB 121 GAGCGGAGCGCTCGCGTCATCTCGTAAATCCACAGCGCGCTCCCGTCTCCACGG 180

QY 181 CCTCACCCCTAGCGATGCGCCACTCCCGCGCTCGTGATCCATGCGCTCACTCCCGTTC 240
DB 181 CCTCACCCCTAGCGATGCGCCACTCCCGCGCTCGTGATCCATGCGCTCACTCCCGTTC 240

QY 241 CTCACACTCTTCGGCGGTCTCGCCAGTGAATAGTGGCGCCCGGAGGGGAGCGCGCG 300
DB 241 CTCACACTCTTCGGCGGTCTCGCCAGTGAATAGTGGCGCCCGGAGGGGAGCGCGCG 300

QY 301 TCGCGGTAGCGACTCCACCGCGCATTTGGCGGTGGGTGGCGCGCGCGCGGAATGG 360
DB 301 TCGCGGTAGCGACTCCACCGCGCATTTGGCGGTGGGTGGCGCGCGCGCGGAATGG 360

QY 361 AGCAGCGCCCGGGAGGCCCGCCAGCCAGTGAAGGAGAGCGCTGTAATGTCGAAAT 420
DB 361 AGCAGCGCCCGGGAGGCCCGCCAGCCAGTGAAGGAGAGCGCTGTAATGTCGAAAT 420

QY 421 TGGGAAATCGACTAATATTTTATGGCAATTTGCTTGATTTGACAAATCTGATAGACAA 480
DB 421 TGGGAAATCGACTAATATTTTATGGCAATTTGCTTGATTTGACAAATCTGATAGACAA 480

QY 481 ATTGCTGGGACAAAAGGCTGTCTGTATGGATAACAGGACTCAGTGGTTTCAGGGAAAAG 540
DB 481 ATTGCTGGGACAAAAGGCTGTCTGTATGGATAACAGGACTCAGTGGTTTCAGGGAAAAG 540

QY 541 TACTCTTGCAATGCACTGAGTCTGAGTGTGATTCAGAGGCCACCTCACTATGTACT 600
DB 541 TACTCTTGCAATGCACTGAGTCTGAGTGTGATTCAGAGGCCACCTCACTATGTACT 600

QY 601 TGATGGTGAACCTCAGACATGGCTTAATAGAGATTTAAGCTTTAAGGCGAAGACCG 660
DB 601 TGATGGTGAACCTCAGACATGGCTTAATAGAGATTTAAGCTTTAAGGCGAAGACCG 660

QY 661 TGCAGAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTGTGATGCTGGTGCAT 720
DB 661 TGCAGAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTGTGATGCTGGTGCAT 720

QY 721 ATGCAATGCTAGCTGATATCTCCATACAGGAGAGATCGTGATGCAATGCCGCTCTACT 780
DB 721 ATGCAATGCTAGCTGATATCTCCATACAGGAGAGATCGTGATGCAATGCCGCTCTACT 780

QY 781 TCCACATCTCACTTATTGAAGTATTTATTCATTTGCCCTTAAATAATTTGGAAGCTCG 840
DB 781 TCCACATCTCACTTATTGAAGTATTTATTCATTTGCCCTTAAATAATTTGGAAGCTCG 840

QY 841 TGATCTTAAAGCCCTATACAGCTTGCACGTCACAGGAAGATTAAAGCTTTCCTGGAAAT 900
DB 841 TGATCTTAAAGCCCTATACAGCTTGCACGTCACAGGAAGATTAAAGCTTTCCTGGAAAT 900

QY 901 TGATGATCCATACGAAACCAACCAATTAATGGTGAATAGTAATTAAGATGAAAGATGAGGA 960
DB 901 TGATGATCCATACGAAACCAACCAATTAATGGTGAATAGTAATTAAGATGAAAGATGAGGA 960

QY 961 ATGCGCTTCAACCCAAAGCAATGGCCAGCAAGTTCTATGCTACCTTGAAGAAAACGGATA 1020
DB 961 ATGCGCTTCAACCCAAAGCAATGGCCAAAGTTCTATGCTACCTTGAAGAAAACGGATA 1020

QY 1021 TTTGCAAGCTTAGTATATATATTTTGAAGAACTGATCTGATTTCTGTGTCTCCATTAAT 1080
DB 1021 TTTGCAAGCTTAGTATATATATTTTGAAGAACTGATCTGATTTCTGTGTCTCCATTAAT 1080

QY 1081 TGTGGACACAATAAGATCTGTTGTTGGTTCATCAATGAATAAAAGGCATCAACATGTAGGAAG 1140
DB 1081 TGTGGACACAATAAGATCTGTTGTTGGTTCATCAATGAATAAAAGGCATCAACATGTAGGAAG 1140

QY 1141 TAAACAGAGGTACCGTTTCATTTCAGAAACGGATATGGATTCATTCGTTTAAAAAAGAAAA 1200
DB 1141 TAAACAGAGGTACCGTTTCATTTCAGAAACGGATATGGATTCATTCGTTTAAAAAAGAAAA 1200

QY 1201 AAAAAAAGAAAAA 1217
DB 1201 AAAAAAAGAAAAA 1217

RESULT 2
AAZ50163
ID AAZ50163 standard; cDNA; 928 BP.
XX AAZ50163;
AC AAZ50163;
XX
DT 04-MAY-2000 (first entry)
XX
DE Wheat Adenylsulphate kinase-1 cDNA clone.
XX
KW Adenylsulphate kinase; Adenosine-5-phosphosulphate kinase; APS kinase;
KW 3'-phospho-adenosine-5-phosphosulphate; PAPS; sulphate assimilation;
KW wheat; clone wrl.pk0101.e2; transgenic plant; screen; antibody; ss.
XX
OS Triticum aestivum.
XX
FH Key Location/Qualifiers
FT CDS 3..743
FT /*tag= a
FT /product= "Wheat APS kinase-1"
FT /note= "Derived from clone wrl.pk0101.e2"
XX
FN WO200004165-A1.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015809.
XX
PR 14-JUL-1998; 98US-0092833P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Falco SC, Allen SM, Anderson SL;
XX
DR WPI; 2000-182430/16.
XX P-PSDB; AAY44792.
XX
PT New nucleic acid molecule and chimeric gene encoding an adenosine-5'
PT phosphosulfate kinase, useful for altering expression of sulfate
PT assimilation protein in plants.
XX
PS Claim 3; Page 34-35; 42pp; English.
XX

CC have a high nutritional value with reduced apical dominance or dwarfism,
CC early flowering or altered metabolic pathways. This sequence represents a
CC plant nucleic acid of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 378 BP; 75 A; 125 C; 89 G; 89 T; 0 U; 0 Other;

Query Match 28.8%; Score 350.2; DB 12; Length 378;
Best Local Similarity 98.7%; Pred. No. 2.3e-78;
Matches 374; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

Qy 245 CACTCTTCGCGGGTCTCCAGTGAATGGGCGCGCGAGGGGAGGGCGCGGTGCG 304

Db 378 CACTCTTCGCGGGTCTCCAGTGAATGGGCGCGCGAGGGGAGGGCGCGGTGCG 319

Qy 305 CGTAGCAGTGCACGGGGGATGGGCGGTGGTGGCGCGCGCGCGGAGGATGGACA 364

Db 318 CGTAGCAGTGCACGGGGGATGGGCGGTGGTGGCGCGCGCGCGGAGGATGGACA 259

Qy 365 GCGCC- CGGGGAGGCGCCGACAGCCAGTGAAGGAGAGCCTGTAATGTCGAACATTGG 423

Db 258 GCGCCAGGGGATGCCCCGACAGCCAGTGAAGGAGAGCCTGTAATGTCGAACATTGG 199

Qy 424 GAAATCGACTAATATTTATGCGCAATTCGTTGATGGACAATCTGATAGACAGAAAT 483

Db 198 GAAATCGACTAATATTTATGCGCAATTCGTTGATGGACAATCTGATAGACAGAAAT 139

Qy 484 GCTGGGCAAAAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 543

Db 138 GCTGGGCAAAAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 79

Qy 544 TCTTGATGTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 603

Db 78 TCTTGATGTCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 20

Qy 604 TGGTGACACCTCAGACAT 622

Db 19 TGGTGACACCTCAGACAT 1

RESULT 4

AAZ50162

ID AAZ50162 standard; cDNA; 936 BP.

XX AAZ50162;

XX 04-MAY-2000 (first entry)

XX Soybean Adenylsulphate kinase cDNA clone.

XX Adenylsulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;

XX 3'-Phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;

XX soybean; clone sdp2c.pk013.all; transgenic plant; screen; antibody; ss.

XX Glycine max.

XX Key Location/Qualifiers

XX 1..936

XX /tag= a

XX /product= "Soybean APS kinase"

XX /note= "Derived from clone sdp2c.pk013.all"

XX WO200004165-A1.

XX 27-JAN-2000.

XX 13-JUL-1999; 99WO-US015809.

XX 14-JUL-1998; 98US-0092833P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX

PI Falco SC, Allen SM, Anderson SL;

XX WPI; 2000-182430/16.

DR P-PSDB; AAY44791.

XX New nucleic acid molecule and chimeric gene encoding an adenosine-5'

PT phosphosulphate kinase, useful for altering expression of sulfate

PT assimilation protein in plants.

XX Claim 3; Page 33; 42pp; English.

XX The present sequence is a cDNA encoding soybean adenylsulphate kinase

CC (APS kinase), also known as adenosine-5' phosphosulphate kinase. This is

CC obtained from clone sdp2c.pk013.all, derived from soybean developing

CC pods, sdp2c cDNA library. APS kinase is a sulphate assimilation protein,

CC that catalyses the conversion of adenosine-5' phosphosulphate (APS) to 3'

CC -Phospho-adenosine-5' phosphosulphate (PAPS). This sequence is used as a

CC probe and primer to identify, obtain and synthesise sulphate assimilation

CC proteins from other plants. It is also used to produce transgenic plants,

CC that are useful for altering the expression levels of a sulphate

CC assimilation protein. The APS kinase peptides are useful for producing

CC antibodies, that are used to screen and isolate cDNA clones

XX Sequence 936 BP; 266 A; 206 C; 240 G; 224 T; 0 U; 0 Other;

Query Match 27.8%; Score 338.8; DB 3; Length 936;

Best Local Similarity 71.5%; Pred. No. 2.6e-75;

Matches 445; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

Qy 410 ATGTGCAACATTTGGAAATCGACTAATATTTATGGACAAATTTGTTGACAAATCT 469

Db 313 ATGTCAAATGTTGGAAATCGACAAACATTTATGGCATGATCTTCCAAATTCAGAAACA 372

Qy 470 GATAGACAGAAATTTGTTGGGACAAAAGGCTGTGTCGTATGGATAACAGGACTCAGTGGT 529

Db 373 GATAGACAGAGCTGCTTCAGCAACAGGCTGTATATGGCTAACCTGGCTCAGCGGA 432

Qy 530 TCAGGAAAAGTACTCTTTGCAATGTGCACTGAGTCTGAGTTGCAATTCAGAGGCACTC 589

Db 433 TCAGGAAAAGCACTATTGTCATGTCCTGAGTCAAAAGCTTGCATCCAAAGGAAAACCTG 492

Qy 590 ACGTATGTAATTTGATGGTGACAACTCAGACATGGCTTAATAGATTTTAAAGCTTTAAG 649

Db 493 TCTTACATCTTGTGTTGACAAATTTGCGCATGGTCTTAAACCCAGGATCTTAGTTTGA 552

Qy 650 GCAGAAGACCTGCAGAAAATATACGAAGAGTTGGTGAAGTGCGAAAAGCTTTTCTGCTGAT 709

Db 553 GCAGAAGATCTGTTGAAAACATTTAGAAGGATTTGGTGAAGTGCGAAAAGCTTTTTCAGAT 612

Qy 710 GCTGGTGTATATGATTTGCTAGCTTATATCTCCATACAGAGAGATCGTGATGATGCG 769

Db 613 GCTGGTGTATTTGTCATCACTAGTTTAAATATCACCATACCAAAAGGATAGAGATGATGCG 672

Qy 770 CGTGCTCTACTTCCACATCTTAACCTTTATGCAAGTATTATTGATTTGCCCTTAAATTT 829

Db 673 AGAGCACTACTTTTCAAAAGGAGATTTTATGAGGTTTTCATAGATGTTTCCACTACATG 732

Qy 830 TGTGAAGCTCGTGTATCTTAAAGGCGCTATACAAAGCTTGCACCTTACAGGAAAGATTAAGGT 889

Db 733 TGTGAAGCTAGGACCCCAAGGAGCTCTCAAGCTTGTCTGAGCTGGAAGATCAAGGT 792

Qy 890 TTCACTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 949

Db 793 TTCACTGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 852

Qy 950 AAAGATGAGGATGCGCTTCCACCAAGCAATGGCCAAAGCAAGTCTTATGCTTACCTTGA 1009

Db 853 AAAGGAGTGAATGTAAGTCTCCCAAGTATGGCTGAAGAGTGAATATCTTACTTGGAG 912

Qy 1010 GAAACCGGATATTTGCAAGCTT 1031

Db 913 GAGAACGGATACCTGCGGCTT 934


```
RESULT 5
AAC44691
ID AAC44691 standard; DNA; 627 BP.
XX
AC AAC44691;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 43768.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0134941P.
PR 21-MAY-1999; 99US-0135124P.
PR 24-MAY-1999; 99US-0135353P.
PR 25-MAY-1999; 99US-0135629P.
PR 27-MAY-1999; 99US-0136021P.
PR 28-MAY-1999; 99US-0136392P.
PR 01-JUN-1999; 99US-0137822P.
PR 03-JUN-1999; 99US-0137222P.
PR 04-JUN-1999; 99US-0137528P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
```

```
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149802P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159328P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162143P.

Query Match 26.3%; Score 320.4; DB 3; Length 627;
Best Local Similarity 69.5%; Pred. No. 9.9e-71;
Matches 435; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 410 ATGTGCACTTGGGAATCGACTAATATTTTATGCGCAATTTGCTTGTATGGCAATCT 469
DB 1 ATGTGCACTTGGGAATTTCAACGACATATTTTGGCAGAAATCCCATTTGGGAAACT 60

QY 470 GATAGACAGAAATTCGTGGGACAAAAGGCTGTGCTATGGAATAACAGGACTCAGTGGT 529
DB 61 GAAAGGCAAGTTGTAAATCAGAAGGGTTGTGTGGTGTGATCACAGGGCTCAGTGGC 120

QY 530 TCAGGAAAGTACTTTCATGTGCATGAGTGTGATTCATTCAGAGGCCACTC 589
DB 121 TCAGGAAAGACCGTTAGCTTGTCTGTAGTAGAGCTGAACAAACCGGGAAGCTA 180
```

```
QY 590 ACGTATGTACTTGTGACAACTCAGACATGCGCTAAATAGAGATTTTAAGCTTTAAG 649
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
181 TCATATATTTCTTGATGGTGACATCTTCGTCACTGTTTGAACAAGATCTTGGTTTCAAG 240

QY 650 GCAGAAGACCGTGCAGAAAATATACGAAGAGTTTGGTGAAGTGGCAAGCTTTTGGCTGAT 709
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
241 GCAGAGGATAGAGTGGAAAATATACGAGGGTCCGAGAAGTAGCCAAACTCTTTGGCGAT 300

QY 710 GCTGGTGCATATGCATTGCTAGCTTGATATCTCCATACAGGAGAGATCGTGATGCATGC 769
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
301 GCTGGTTTAACTCTGTAATGCCAGCCTCATATCCCGCTATAGANAAGACCGTGACGCCCTGC 360

QY 770 CGTCTCTACTTCCACATTTCTAACTTTTATTGAAGTATTATTGATTTGCCCTTAAAAATT 829
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
361 AGGGAATGATACAGAATTGATCTTTTATTGAGGTTTTCATGAACATGTCTCTGCAATTG 420

QY 830 TGTGAAGCTCGTGATCTCTAAAGGCTATACAAAGCTTGCACGTACAGGAAAGATTAAAGGT 889
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
421 TGTGAAGCAAGGACCCCTTAAAGGCTATACAAAGCTTGCACGTGCAGGAAAGATCAAGGT 480

QY 890 TTCACTGGAATTGATGATCCATACGAACCAATTAATGCTGAGATAGTAATTAAGATG 949
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
481 TTCACAGGAATAGATGATCCGATATGAGTCTCCCTTGAACCTGTGAGATAGAGTTGAAAGAG 540

QY 950 AAAGATGAGGAATGCCCTTTCACCCAAAGCAATGGCCAAGCAAGTTCTATGCTACCTTGAA 1009
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
541 AAAGAGGAGAGTGTCTTCCCTGTAGTATGCTGAGAGTGAATCTCTTATCTAGAA 600

QY 1010 GAAACCGGATATTTGCAAGCTTAGTA 1035
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
601 GACAAAGGTTCTTCTCAAAACGAGTA 626

RESULT 6
AAZ50159
ID AAZ50159 standard; cDNA; 890 BP.
XX
AC AAZ50159;
XX
DT 04-MAY-2000 (first entry)
XX
DE Corn Adenylisulphate kinase-1 cDNA clone.
XX
KW Adenylisulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
KW 3'-Phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
KW corn; clone cen3n.pk0088.b10; transgenic plant; screen; antibody; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
CDS 3..677
FT /*tag= a
FT /product= "Corn APS kinase-1"
FT /note= "Derived from clone cen3n.pk0088.b10"
XX
PN WO200004165-A1.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015809.
XX
PR 14-JUL-1998; 98US-0092833P.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Falco SC, Allen SM, Anderson SL;
XX
DR WPI; 2000-182430/16.
XX P-PSDB; AAY44788.
XX
PT New nucleic acid molecule and chimeric gene encoding an adenosine-5' phosphosulfate kinase, useful for altering expression of sulfate
```

PT assimilation protein in plants.

XX Claim 3; Page 29; 42pp; English.

XX The present sequence is a cDNA encoding corn adenylsulphate kinase (APS
XX kinase), also known as adenosine-5' phosphosulphate kinase. This is
XX obtained from clone cen3n.pk0088.b10, derived from corn endosperm, cen3n
XX cDNA library. APS kinase is a sulphate assimilation protein, that
XX catalyses the conversion of adenosine-5' phosphosulphate (APS) to 3'-
XX phospho-adenosine-5' phosphosulphate (PAPS). This sequence is used as
XX probes and primers to identify, obtain and synthesise sulphate
XX assimilation proteins from other plants. It is also used to produce
XX transgenic plants, that are useful for altering the expression levels of
XX a sulphate assimilation protein. The APS kinase peptides are useful for
XX producing antibodies, that are used to screen and isolate cDNA clones
SQ Sequence 890 BP; 232 A; 239 C; 262 G; 157 T; 0 U; 0 Other;

Query Match 26.0%; Score 316.2; DB 3; Length 890;

Best Local Similarity 66.9%; Pred. No. 1.3e-69;

Matches 450; Conservative 0; Mismatches 223; Indels 0; Gaps 0;

QY 361 AGCAGCGCCGGGGGCCCCGACAGCCAGTGAAGGAGAGCCTGTAATGTCGAACAT 420

Db 5 AGCGCGCGCGCGCTGCGAGGATCAGCAGCAGCAGCGCGCTGGTGACCTGCACCGT 64

QY 421 TGGGAATCGCATTAATATTTATGCAATATGCTTGTGACAACTCTGATAGACAGAA 480

Db 65 CGGGAATCGCAACATCTCTGTGATGATGCGCATCGGGCAGAGGAGCGACAGGG 124

QY 481 ATTGCTGGGAAAAAGGCTGTGCTATGTAATACAGGACTCAGTGGTTTCAGGAAAAAG 540

Db 125 TCTGCTGAACCAAGAGGGCTGCTGTGTGATCATCTGSCCTAAGCGGTTTCAGGGAAG 184

QY 541 TACTCTTGATGTGACTGAGTCGTGATGTCATTCGAGGCGCACCTCAGTATGTAAT 600

Db 185 CACGCTCGCGCTGCGCTGAGCGCGAGCTGACGCGAGGCGCACCTCAGTATGTAAT 244

QY 601 TGATGTGACAACTTCAGACATGGCTAAATAGAGATTAAAGCTTAAAGGAGAGACCG 660

Db 245 CGACGGGCAACCTTCAGCAGCGGCTGAAAGGACCTCAGCTTCGGAGCAGAGACCG 304

QY 661 TCGAAGAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTGTCTGATGCTGGTGCAT 720

Db 305 CGCCGAGAACATCCGACAGATAGGGAAGTAGCGAAGCTGTTGCGCGAGCTGGCTCGT 364

QY 721 ATGCAATGCTAGCTTGATATCTCCATACAGGAGATGATGATGATGATGATGATGAT 780

Db 365 CTGCATCGCCAGCTCATATCGCCCTACAGAAAGCGACCGCGCTGTCGCGATCTGCT 424

QY 781 TCCACATTCCTAATTTATGAAATATTTATTCATTTGCCCTTAAATTTTGAAGCTCG 840

Db 425 GCCAAGCACTCGTTTATCGAGTGTCTTGAAGCTGCGCTTCAAGTGTGCGAAGCCAG 484

QY 841 TGATCCTTAAAGCCCTATCAAGCTTGCAGTACAGGAAAGATTAAGGTTTCACTGGAAT 900

Db 485 GGACCCCAAGGCTCTCAAGCTCGACGCGCGCAAAATCAAGGGTTCAACCGCAT 544

QY 901 TGATGATCCTACGAAACCAATTAATGCTGAGATAGTAAATTAAGATGAAGATGAGA 960

Db 545 CGAGATCTTACGAACCGCGCTCGGAGCTGTGAGATAGTATCCAGTGAAGTCCGCGA 604

QY 961 ATGCCCTTACCACAAAGCAATGGCCAGCAAGTCTATGCTACCTTGAAGAAACGGATA 1020

Db 605 CTGCCCTTCCCTGATCGATGCTGCTGGTGCACCTGTGCTGCTGATGAGCAATGGTTT 664

QY 1021 TTTCGAAGCTTAG 1033

Db 665 CTTCCAGGACTAG 677

RESULT 7
AAC35266

ID AAC35266 standard; DNA; 1068 BP.

XX AAC35266;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 9570.

XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-01231180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 08-APR-1999; 99US-0128714P.

XX 16-APR-1999; 99US-0129845P.

XX 19-APR-1999; 99US-0130077P.

XX 21-APR-1999; 99US-0130449P.

XX 23-APR-1999; 99US-0130510P.

XX 23-APR-1999; 99US-0130891P.

XX 28-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 30-APR-1999; 99US-0132407P.

XX 04-MAY-1999; 99US-0132484P.

XX 05-MAY-1999; 99US-0132485P.

XX 06-MAY-1999; 99US-0132486P.

XX 06-MAY-1999; 99US-0132487P.

XX 07-MAY-1999; 99US-0132863P.

XX 11-MAY-1999; 99US-0134256P.

XX 14-MAY-1999; 99US-0134218P.

XX 14-MAY-1999; 99US-0134219P.

XX 14-MAY-1999; 99US-0134221P.

XX 14-MAY-1999; 99US-0134370P.

XX 18-MAY-1999; 99US-0134768P.

XX 19-MAY-1999; 99US-0134941P.

XX 20-MAY-1999; 99US-0135124P.

XX 21-MAY-1999; 99US-0135353P.

XX 24-MAY-1999; 99US-0135629P.

XX 25-MAY-1999; 99US-0136021P.

XX 27-MAY-1999; 99US-0136392P.

XX 28-MAY-1999; 99US-0136782P.

XX 01-JUN-1999; 99US-0137222P.

XX 03-JUN-1999; 99US-0137528P.

XX 04-JUN-1999; 99US-0137502P.

XX 07-JUN-1999; 99US-0137724P.

XX 08-JUN-1999; 99US-0138094P.

XX 10-JUN-1999; 99US-0138540P.

XX 10-JUN-1999; 99US-0138847P.

XX 14-JUN-1999; 99US-0139119P.

XX 16-JUN-1999; 99US-0139452P.

XX 16-JUN-1999; 99US-0139453P.

XX 17-JUN-1999; 99US-0139492P.

XX 18-JUN-1999; 99US-0139454P.

XX 18-JUN-1999; 99US-0139455P.

XX 18-JUN-1999; 99US-0139456P.

XX 18-JUN-1999; 99US-0139457P.

XX 18-JUN-1999; 99US-0139458P.

XX 18-JUN-1999; 99US-0139459P.

XX 18-JUN-1999; 99US-0139460P.

Db 488 GTTATATCTTGATGTGATGAATAATGTTAGGCATGGCTTAAACCGTGATCTTAGCTTTAAAG 547
QY 651 CAGAAGACCGTGCAGAAATAATACGAAGAGTGGTGAAGTGGCAAGCTTTTGTCTGATG 710
Db 548 CTGAGGATCGTCAGAGATATTCGTAGAGTGGAGAGGTTGCTAAGCTTTTTCGGATG 607
QY 711 CTGGTGTATATGATCTGATGCTTAGCTTGCATATCTCCATACAGGAGATCTGATGATGCC 770
Db 608 CTGGAATAATCTGCATTCGAGTTTGCATATCTCTTATAGAACAGATAGGAGCGCTTGTG 667
QY 771 GTGCTCTACTTCCACATCTTAACCTTTATTTGAAGTATTTATTTGCTGCTTCCCTTAAATTT 830
Db 668 GAAGTTTGTCTCCCGAGGAGATTTTGTGAGTGTTCATGATGATCGCTTAGTGTTT 727
QY 831 GTGAAGCTCGTATCTCAAGGCTATACAGCTTCACGATACAGGAAGATTTAAAGTT 890
Db 728 GCGAGCGGAGGATCCAAAGGTTCTTTACAGCTTCTCTGCGAGGAAGATCAAGTT 787
QY 891 TCACTGGAATTGATGATCCATACGAACCAACCAATTAAGTGTGATAGTAATTAAGATGA 950
Db 788 TTACCGGGATCGATGACCTTACGAGCCACCATTTGAATGCGAG-----ATTCTCTAG 841
QY 951 AGATGAGGAATGCCCTTACCCAAAGCAATGCCCAAGCAAGTTCTATGCTACCTTGAAG 1010
Db 842 GAGTGAAGGAGGAACCTTCTCTATCGAAATGCGGAAAGGTCGTCGGATACTTAGATA 901
QY 1011 AAAACGATATTGCAAGCTTAGTATATGAT 1042
Db 902 ACAGGGTTATCTTCAAGCATACATACCTTCT 933

RESULT 8
AAC36588
ID AAC36588 standard; DNA; 1175 BP.
XX AC AAC36588;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 14353.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 22-JUL-1999; 99US-0145089P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.

XX Hybridisation assay; Genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KN promoter; termination sequence, ss.
XX Arabidopsis thaliana.
OS EP1033405-A2.
XX 06-SEP-2000.
PD 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-01218025P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0123788P.
PR 29-MAR-1999; 99US-0126264P.
PR 01-APR-1999; 99US-0126785P.
PR 06-APR-1999; 99US-0127462P.
PR 08-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 19-APR-1999; 99US-0129845P.
PR 21-APR-1999; 99US-0130077P.
PR 23-APR-1999; 99US-0130449P.
PR 28-APR-1999; 99US-0130510P.
PR 30-APR-1999; 99US-0130891P.
PR 04-MAY-1999; 99US-0131449P.
PR 05-MAY-1999; 99US-0132048P.
PR 06-MAY-1999; 99US-0132407P.
PR 07-MAY-1999; 99US-0132486P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 28-JUN-1999; 99US-0140695P.
PR 29-JUN-1999; 99US-0140823P.
PR 30-JUN-1999; 99US-0140991P.
PR 01-JUL-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.

PR	31-AUG-1999;	99US-0151438P.	
PR	01-SEP-1999;	99US-0151930P.	
PR	07-SEP-1999;	99US-0152363P.	
PR	10-SEP-1999;	99US-0153070P.	
PR	13-SEP-1999;	99US-0153758P.	
PR	15-SEP-1999;	99US-0154018P.	
PR	16-SEP-1999;	99US-0154039P.	
PR	20-SEP-1999;	99US-0154779P.	
PR	22-SEP-1999;	99US-0155139P.	
PR	23-SEP-1999;	99US-0155486P.	
PR	24-SEP-1999;	99US-0155659P.	
PR	28-SEP-1999;	99US-0156458P.	
PR	29-SEP-1999;	99US-0156596P.	
PR	04-OCT-1999;	99US-0157117P.	
PR	05-OCT-1999;	99US-0157753P.	
PR	06-OCT-1999;	99US-0157865P.	
PR	07-OCT-1999;	99US-0158029P.	
PR	08-OCT-1999;	99US-0158232P.	
PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161406P.	
PR	26-OCT-1999;	99US-0161359P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161992P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match 24.4%; Score 296.4; DB 3; Length 1183;			
Best Local Similarity 68.0%; Pred. No. 1.6e-64;			
Matches 430; Conservative 0; Mismatches 196; Indels 6; Gaps 1;			
QY	411	TGTCGAACATTTGGGAAATCGACTAATATTTATGGCACAATTCGTTGATTTGACAAATCTG 470	
Db	309	TGTCCTACGGTCGAAACTCGACAAATATAAAGTGGCATGAATGTTCTCTTTGAGAAAAGTTG 368	
QY	471	ATAGACAGAAATTCCTGGGCAAAAAGGCTGTGCTGATCGATACAGGACTCAGTGGTT 530	
Db	369	ATAGACAGAGATGCTTGTATCAGAAGAGATGTGATTTGGGTCAACCGGTCTTAGTGGTT 428	
QY	531	CAGGGAAGAGTACTCTTCGATGTGCACGTGAGTCGTGAGTTGCATTCGAGAGGCCACCTCA 590	
Db	429	CAGGGAAGAGTACTTTGGCTTTGCTTTGATCAGATGTTGTATCAAAAGGGGAAGCTTT 488	
QY	591	CGTATGTACTTCGATGGTGACAACTCAGACATGGCCCTAAATAGAGATTTAAGCTTTAAG 650	
Db	489	GTATATATTCTTGATGGTGATAATGTTAGGCATGGCTTAAACCGGTGATCTTAGCTTTAAG 548	
QY	651	CAGAAGACGTCGAGAAATATACGAGAGTTGCTGAAGTGGCAAGACTTTTGTCTGATG 710	
Db	549	CTGAGGATCGTCGAGAAATATTCGTAGAGTTGGAGAGGTTGCTAAGCTTTTTCGCGATG 608	
QY	711	CTGGTGTCTATATGCATTTGCTAGCTTGATATCTCCATACAGGAGAGATCGTGATGCATGCC 770	

Db	609	CTGGAATAATCTGCATTTGCGAGTTTGATATCTCTTATAGAACAGATAGGACGCTTGTG 668
QY	771	GTGCTCTACTTCCACATTTCTAACTTTATTTGAAGTATTTATTTGTTGCCCTCAAAAATTT 830
Db	669	GAAGTTTGCTCCCGAGGAGATTTTGTGAGGTTCATCGGATGTACCGTTAGTGTGTTT 728
QY	831	GTGAAGCTCGTGATCCTAAAGGCCCTATACAAGCTTGCAAGCTTGACAGGAAAGATTAAAGTT 890
Db	729	GCGAGGCGAGGATCCAAAGGCTCTTTACAAGCTTGCTCGTGAGGAAAGATCAAGGTT 788
QY	891	TCACCTGGAATGTGATCCATACGAACCAACCAATTAATGTGAGATAGTATTAAGATGA 950
Db	789	TTACCGGATCGATGACCTTTACGAGCCACCAITGAACTGCGAG-----ATTTCCTCTAG 842
QY	951	AAGATGAGGATGCCCTTACCCAAAGCAATGGCCAAAGAGTTCCTATGCTTACCTTGAAG 1010
Db	843	GACGTGAAGGAGGAATCTCTCTATCGAATGCGGAAAGGTCGTCGGATCTTAGATA 902
QY	1011	AAAAACGGATATTTGCAAGCTTAGTATATGTAT 1042
Db	903	ACAAGGGTTATCTTCAAGCATACATACTTCT 934
RESULT 10		
ABZ12905		
ID	ABZ12905	standard; DNA; 831 BP.
AC	ABZ12905;	
XX	21-JAN-2003	(first entry)
DE	Arabidopsis thaliana	stress regulated gene SEQ ID NO 710.
DE	Arabidopsis thaliana	plant; gene; stress; transgenic; ds.
KW	Arabidopsis thaliana.	
OS	WO200216655-A2.	
PN	28-FEB-2002.	
XX	24-AUG-2001;	2001WO-US026685.
XX	24-AUG-2000;	2000US-0227866P.
PR	26-JAN-2001;	2001US-0264647P.
PR	22-JUN-2001;	2001US-0300111P.
XX	(SCRI)	SCRIPPS RES INST.
PA	(SYGN)	SYNGENTA PARTICIPATIONS AG.
XX	Harper JF, Kreps J, Wang X, Zhu T;	
XX	WPI;	2002-304127/34.
DR	Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses.	
PT	Claim 144; SEQ ID NO 710; 577pp + Sequence Listing; English.	
XX	The invention relates to identifying a stress condition to which a plant cell has been exposed, comprising: (a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and (b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response; The method is useful in the production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stress. The present sequence is that of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used in methods of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office	

Db 515 CTGGAATAATCTGCATTGCGAGTTTGATATCTCTTATAGAACAGATAGGACGCTTGTC 574
Qy 771 GTGCTCTACTCCACATCTTAACCTTATTGAAGTATTTATGATTTGCCCTAAAAATTT 830
Db 575 GAAGTTTGTCTCCCGAGGAGATTTTGTGAGGTTCATGGATGTACCGCTTAGTGTT 634
Qy 831 GTGAAGCTCGTATCTCTAAAGCCCTATACAAAGCTTGCAAGCTTGACAGAAAGATTAAGGTT 890
Db 635 GCGAGGCGAGGATCCAAAGGCTTTACAAGCTTGCTGTCGAGAAAGATCAAGGTT 694
Qy 891 TCACCTGGAATTGATGATCATACGACCAACCAATTAAGTGGTGGATAGTAATTAAGATGA 950
Db 695 TTACCGGGATCGATGACCTTACGAGCCACCAATGAACCTGCCAG-----ATTCTCTAG 748
Qy 951 AGATGAGGAAGTCCCTTACCCAAAGCAATGCGCAAGCAAGTTCATGCTACCTTGAAG 1010
Db 749 GACGTGAAGGAGAACTTCTCTATCGAAATGGCGAAAGGTCGTGCGATACCTTAGATA 808
Qy 1011 AAAACGGATATTTGCAAGCTTA 1032
Db 809 ACAAGGTTATCTTCAAGCATA 830

RESULT 12

ADN74374
ID ADN74374 standard; cDNA; 831 BP.

AC ADN74374;

XX 15-JUL-2004 (first entry)

XX Thale cress cDNA repressed in E2Fa/Dpa expressing plants SeqID 2269.

DE gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;

KW growth regulator; animal feed product; thale cress;

KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.

XX Arabidopsis thaliana.

XX WO2004035798-A2.

XX 29-APR-2004.

XX 20-OCT-2003; 2003WO-BP011658.

XX 18-OCT-2002; 2002EP-00079408.

XX (CROP-) CROPDISEIGN NV.

XX Inze D, De Veylder L, Vlieghe K;

XX WPI: 2004-348466/32.

DR P-PSDB; ADN74375.

XX Altering plant characteristics, useful for producing plants for enzyme or
PT pharmaceutical production comprises modifying in a plant, expression of
PT one or more nucleic acids and/or modifying level or activity of one or
PT more proteins.

XX Claim 1; SEQ ID NO 2269; 134pp; English.

XX This invention relates to a novel method for altering one or more plant
CC characteristics. Specifically, it refers to identifying genes that are up
CC - or down-regulated in transgenic plants overexpressing the heterodimeric
CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
CC alter plant characteristics accordingly. The present invention describes
CC generating transgenic plants for the production of growth regulators,
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
CC the altered plant characteristics are selected from increased yield or
CC biomass, enhanced survival capacity, stress tolerance, plant architecture
CC or physiology, altered endoreduplication, signal transduction, signal
CC transduction, storage lipid mobilization and/or altered photosynthesis,
CC each relative to the corresponding wild type plants. Accordingly, these

CC sequences can also be useful as positive or negative selectable markers
CC during transformation of cells or tissues. The identified genes play a
CC role in a variety of biological processes such as DNA replication, cell
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
CC transcription factors. This polynucleotide sequence is thale cress cDNA
CC repressed 1.3 fold or more in plants overexpressing the E2Fa/Dpa
CC transcription factor, given in an exemplification of the invention.

XX SQ Sequence 831 BP; 224 A; 137 C; 223 G; 247 T; 0 U; 0 Other;

Qy Query Match 24.2%; Score 294.4; DB 12; Length 831;

Db Best Local Similarity 68.3%; Pred. No. 4.3e-64;

Matches 425; Conservative 0; Mismatches 191; Indels 6; Gaps 1;

Qy 411 TGTGCAACATTTGGGAATTCGACTTAATATTTTATGSCACAAATTTGCTTGATTGGACAATCTG 470

Db 215 TGTCTACCGTCGGAACCTCGACAAATATAAGTGGCATGAATGTCTGTTGAGAAAGTTG 274

Qy 471 ATAGACAGAAATTCGTGGGACAAAAGGCTGTGCTATGGATAACAGGACTCAGTGTGTT 530

Db 275 ATAGACAGAGATTGCTTGATCAGAAAGGATGTGATTTGGGTCAACGGTCTTAGTGTGTT 334

Qy 531 CAGGAAAAGTACTCTTCGATGTGCATGTGCATGTGAGTTCGATTCGAGAGCCACCTCA 590

Db 335 CAGGGAAGAGTACTTTGGCTTTGCTTTGAATCAGATGTTGTATCAAAAGGGGAAGCTTT 394

Qy 591 CGTATGTAATTTGATGGTGCAACACCTCAGACATGGCTTAAATAGAGATTTAAAGG 650

Db 395 GTTATATTTCTTGATGGTGATTAATGTTAGGCATGCTTAAACCGTGATCTTAGCTTTAAG 454

Qy 651 CAGAAGCCGTGCAGAAAATATACGAAGATGGTGAAAGTGGCAAGCTTTTGTCTGATG 710

Db 455 CTGAGGATCGTGCAGAGAATATTCGTAGATTTGGAGAGGTTGCTTAAGCTTTTGGCGATG 514

Qy 711 CTGTGTCATATGATCTGCTAGCTTGATATCTCCATACAGGAGATCGTGTGATGATGCC 770

Db 515 CTGGAATTAATCTGCATTCGAGTTTGATATCTCCTTATAGAACAGATAGGACGCTTGTG 574

Qy 771 GTGCTCTACTTCCACATTTCTAACTTTATTGAAGTATTTATTGATTTGCCCTTAAAAATTT 830

Db 575 GAAGTTTGTCTCCCGAGGGAGATTTTGTGAGTGTCTCATGATGTCACGCTTAGTGTT 634

Qy 831 GTGAAGCTCGTGATCTCTAAAGCCCTATACAGCTTTGCACGTCACAGAAAGATTTAAAGTT 890

Db 635 GCGAGGCGAGGGATCCAAAGGGTCTTTACAAAGCTTGCTCGTCGAGAAAGATCAAGGTT 694

Qy 891 TCACCTGGAATTCATCATCATCAACCAACCAATTAATGCTGAGATACTAATTAAGATGA 950

Db 695 TTACCGGATCGATGACCTTACGAGCCACCATTTGAACCTGCGAG-----ATTCTCTAG 748

Qy 951 AAGATGAGGAATGCCCTTTCACCCAAAGCAATGGCCCAAGAGTTCTATGCTACCTTTGAAG 1010

Db 749 GACGTGAAGGAGGAACCTTCTCTATCGAAATGGCGAAAGGTCGTGCGATACTTAGATA 808

Qy 1011 AAAACGGATATTTGCAAGCTTA 1032

Db 809 ACAAGGTTATCTTCAAGCATA 830

RESULT 13

AAC49070

ID AAC49070 standard; DNA; 1130 BP.

XX AAC49070;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 59815.

DE Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway; metabolic pathway;

KW promoter; termination sequence; ss.

XX

OS Arabidopsis thaliana.
XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301433.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.

PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.

PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match

23.7%; Score 288.4; DB 3; Length 1133;

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 9, 2005, 09:35:59 ; Search time 248 Seconds
(without alignments)
8029.632 Million cell updates/sec

Title: US-10-829-432-3
Perfect score: 1217
Sequence: 1 ggcgcgttcattcatca.....aaaaaaaaaaaaaaaaaaaa 1217

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description	
		Match	Length				
1	159.6	13.1	615	3	US-09-134-001C-1680	Sequence 1680, Ap	
2	156.4	12.9	1160	3	US-09-153-310-41	Sequence 41, Appl	
3	153.2	12.6	2014	4	US-09-898-1658-1	Sequence 1, Appl	
4	152.2	12.5	2015	4	US-09-949-016-1925	Sequence 1925, Ap	
5	152	12.5	1845	4	US-09-898-1658-9	Sequence 9, Appl	
6	150.6	12.4	2617	4	US-09-786-240-21	Sequence 21, Appl	
c	7	146	12.0	640681	4	US-09-790-988-1	Sequence 1, Appl
8	139	11.4	1875	4	US-09-949-016-1468	Sequence 1468, Ap	
9	139	11.4	2506	1	US-08-879-561-4	Sequence 4, Appl	
10	132.8	10.9	1851	4	US-09-898-1658-10	Sequence 10, Appl	
11	132.8	10.9	2000	4	US-09-898-1658-2	Sequence 2, Appl	
12	128.4	10.6	681	4	US-09-543-681A-2995	Sequence 2995, Ap	
13	119.4	9.8	651	4	US-09-489-039A-2803	Sequence 2803, Ap	
14	94.2	7.7	89892	4	US-09-949-016-13667	Sequence 13667, A	
15	80.6	6.6	109925	4	US-09-949-016-13210	Sequence 13210, A	
c	16	79.6	6.5	601	4	US-09-949-016-50640	Sequence 50640, A
17	73.2	6.0	4411529	3	US-09-103-840A-1	Sequence 1, Appl	
18	71.6	5.9	4403765	3	US-09-103-840A-2	Sequence 2, Appl	
19	48.4	4.0	1141	4	US-09-806-708B-22	Sequence 22, Appl	
20	48.4	4.0	1959	4	US-09-252-991A-7983	Sequence 7983, Ap	
c	21	48.4	4.0	2010	4	US-09-252-991A-8169	Sequence 8169, Ap
c	22	47.8	3.9	7218	1	US-08-232-463-14	Sequence 14, Appl
23	45.6	3.7	1323	3	US-09-413-452-3	Sequence 3, Appl	
24	45.6	3.7	1323	3	US-09-413-068-3	Sequence 3, Appl	
25	45.6	3.7	1323	3	US-09-403-345A-3	Sequence 3, Appl	
26	45.4	3.7	98844	3	US-09-791-211-10	Sequence 10, Appl	
27	45.4	3.7	143776	4	US-09-949-001-29	Sequence 29, Appl	

28	45.4	3.7	144034	4	US-09-949-001-35	Sequence 35, Appl	
c	29	45	3.7	318	3	US-09-165-264-12	Sequence 12, Appl
30	45	3.7	3116	4	US-09-904-615-43	Sequence 43, Appl	
31	44.8	3.7	30782	4	US-09-949-016-13724	Sequence 13724, A	
c	32	44.6	3.7	12865	4	US-09-902-540-1048	Sequence 1048, Ap
c	33	44.2	3.6	11766	4	US-09-949-016-12531	Sequence 12531, A
c	34	44.2	3.6	11770	4	US-09-949-016-12720	Sequence 12720, A
c	35	44.2	3.6	11770	4	US-09-949-016-13487	Sequence 13487, A
c	36	44.2	3.6	11770	4	US-09-949-016-13488	Sequence 13488, A
37	43.4	3.6	2447	2	US-09-014-969-14	Sequence 14, Appl	
38	43.2	3.5	1131	4	US-09-902-540-4972	Sequence 4972, Ap	
c	39	43.2	3.5	27219	4	US-09-902-540-1244	Sequence 1244, Ap
40	43	3.5	601	4	US-09-902-540-1244	Sequence 1244, Ap	
41	43	3.5	601	4	US-09-949-016-138313	Sequence 138313, A	
42	43	3.5	1793	4	US-09-949-016-138314	Sequence 138314, A	
43	43	3.5	3875	4	US-09-902-540-6861	Sequence 6861, Ap	
44	43	3.5	8671	4	US-09-902-540-579	Sequence 579, App	
c	45	43	3.5	46603	4	US-09-949-016-15455	Sequence 15455, A
46	43	3.5	46603	4	US-09-949-016-15636	Sequence 15636, A	

ALIGNMENTS

RESULT 1

US-09-134-001C-1680
; Sequence 1680, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1680
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1680

Query Match	13.1%;	Score	159.6;	DB	3;	Length	615;
Best Local Similarity	55.9%;	Pred. No.	1.9e-33;				
Matches	303;	Conservative	0;	Mismatches	239;	Indels	0;
Gaps	0;						
QY	416	AACATTGGGAATCGACTAATATTTATGCGACAAATTCCTTGATTGGACAATCTGATAGA	475				
Db	13	AAATGAGTGAATCAATCATATTACATGGCATGCTCAGAAATTACGAAGAACAAAGA	72				
QY	476	CAGAAATTCCTGGGACAAAAGGCTGCTCGTATGATGAACAGGACTCAGTGGTTCAGGG	535				
Db	73	CAGCATAAAAATGGACACAAAAGTCTGTATATGTTTACAGGCTATCTGGGTTCAGGG	132				
QY	536	AAAGTACTCTTGATGTGCAATGAGTGTGCTGATGTCATTGCAGAGGCCACCTCAGGTAT	595				
Db	133	AAATCAACGGTCTCTGTTGCCTTAGAAAAAGAAATTTATTAATGAAGGAAAAACAACATAC	192				
QY	596	GTACTTGTGTGACAACTCAGACATCGCCTTAATAGAGATTTTAAGCTTTTAAGGCAGAA	655				
Db	193	CGTTAGATGGTGATATGTCGTCATGCTTCTTAATAAAATTTAGGGTTAGTCTGTA	252				
QY	656	GACCGTCAGAAAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTTCGTATGCTGGT	715				
Db	253	GATCGTTCAGAAAAATATTCGACGCAATTCGAGAAGTAGCAAAATTAATGTTAGACGCTGA	312				
QY	716	GTCATATGCATGTAGCTTGATATCTCCATACAGGAGATCGTGATCGCATCGCGTCT	775				
Db	313	GCCTTAACAGTTACTGCTTTTATCTCCCATATAAAGAAGACAGAGAAGGTGTTAGACGA	372				

Db 423 GCCATTCTTTGAAATATTTGTAGATGCACCTCTAAATATTTGTGAAAGCAGAGACGTAAA 482
QY 850 AGGCTTATACAGCTTGCAGCTACAGGAAGATTAAGGTTTCACTTGAATTTGATGATCC 909
Db 483 AGGCTCTATAAAGGCGCAGAGCTGGGAGATTAAGGATTTACAGGATTTGATTTCTGA 542
QY 910 ATACGAACCAACCAATTAATGGTGAGATAGTAATTAAGATGAA 951
Db 543 TTATGAGAAACCTGAAACCTCTGAGCGTGTCTTAAACCAA 584

RESULT 6
US-09-786-240-21
; Sequence 21, Application US/09786240
; Patent No. 6558935
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: AZIMZAI, Yalda
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
; FILE REFERENCE: PF-0592 PCT
; CURRENT APPLICATION NUMBER: US/09/786,240
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/150,657; unassigned; 09/186,779; unassigned; 60/133,642
; PRIOR FILING DATE: 1998-09-10; 1998-09-10; 1998-11-04; 1998-11-04; 1999-05-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PERL Program
; SEQ ID NO 21
; LENGTH: 2617
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6558935 1420940CB1
US-09-786-240-21

Query Match 12.4%; Score 150.6; DB 4; Length 2617;
Best Local Similarity 59.5%; Pred. No. 1.3e-30;
Matches 275; Conservative 0; Mismatches 181; Indels 6; Gaps 1;
QY 496 AGGCTGTGTCGTATGATTAACAGGACTCAGTGGTTTCAGGGAAGATCTCTTGCATGTGC 555
Db 215 AGGATGTACCGTGTGGCTAACAGGTCTCTCTGGTGTGGAAAAACAACGATAAGTTTTCG 274
QY 556 ACTGAGTCGTGAGTTGCAATTCAGAGGCCACCTCAGCTATGATCTTGTGTTGACAACT 615
Db 275 CCTGGAGGAGTACCTTGTCTCCATGCCATCCCTTGTACTCCCTGGATGGGACAATGT 334
QY 616 CAGACATGCGCTAAATAGAGATTTAAGCTTTAAGGCAGAGACCGTGCAGAAAAATATAG 675
Db 335 CGTCATGSCCTTAACAGAACTCGATCCTCTCTCTGGGACAGAGAGAAAAATATCG 394
QY 676 AAGAGTTGTGAGTGGCAAGCTTTTTCCTGATGCTGTGTGCATATGCAATTTGCTAGCTT 735
Db 395 CCGGATTTGTGAGGTGGCTTAAGCTGTTTTCTGATGCTGGTCTGCAATTTACCAGCTT 454
QY 736 GATATCTCCATACAGAGAGATCGTGATGCATGCCGTGCTCTACTT-----CCACATTC 789
Db 455 TTATTTCTCATTCGCGAAGGATCGTGAGATGCCGCCCAAAATACATGAATCAGAGGGCT 514
QY 790 TAACTTTTATGAAGTATTTTATTTGATTTGCCCTTAAATAATTTGTGAAGCTCGTGATCCTAA 849
Db 515 GCAATTTCTTGAATATTTTGTAGATGCACCTCTAAATATTTGTGAAAGCAGAGACGTAAA 574
QY 850 AGGCTTATACAGCTTGCAGCTACAGGAAGATTAAGGTTTCACTGGAATTTGATGATCC 909
Db 575 AGGCTCTATAAAGGCGCAGAGCTGGGAGATTAAGGATTTACAGGATTTGATTTCTGA 634

QY 910 ATACGAACCAACCAATTAATGGTGAGATAGTAATTAAGATGAA 951
Db 635 TTATGAGAAACCTGAAACCTCTGAGCGTGTCTTAAACCAA 676
RESULT 7
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 12.0%; Score 146; DB 4; Length 640681;
Best Local Similarity 55.0%; Pred. No. 4.9e-28;
Matches 287; Conservative 0; Mismatches 235; Indels 0; Gaps 0;
QY 434 AATATTTTATCGCACAAATTTGCTTGGCAATCTGTATACACAGAAATTCCTGGGACAA 493
Db 457777 AATATTTTGGCAAAACATTCATCAACGCTTAAACCGTGAATAAATAGTGTCTAT 457718
QY 494 AAGGCTGTGCTATGGATAACAGGACTCAGTGGTTTCAGGAAAAAGTACTCTTTGCATGT 553
Db 457717 AATCAATTTGACTATGTTCTCCTGGGCTCTCAGGTTTCAGGAAAAATCAACCATAGCAA 457658
QY 554 GCATGAGTCGTGAGTTGCATTCGACAGGCCACCTCAGCTATGATGTACTTGTGTTGACAA 613
Db 457657 TTTTTCAGAGAAATATTTCTTTAAAAATGSAATTAATAGCTATTTATTAGACGGAGACA 457598
QY 614 CTCGACATGCGCTAAATAGAGATTTAAGCTTTAAGGCAGAGACCGTGCAGAAAAATATA 673
Db 457597 ATTAGTCTGTTTATGTTCTGATTTAAGTTTTAGTTTGGCAGATAGGAATGAAAAACAT 457538
QY 674 CGAAGAGTTGCTGAAGTGGCAAGCTTTTTCCTGATGCTGCTGTCTCATATGCAITTCCTAGC 733
Db 457537 AGACGATTTGGAGAAGTAGTAATAATGATTTACATGCTGCTCTAATAATTTGGTATCA 457478
QY 734 TTGATATCTCCATACAGAGAGATCGTGATGCGATGCGCTCTACTTCCCAATTTCTTAAC 793
Db 457477 GTTATTTCCCTATAGAAATCAAAGGGAATTTGGTACGCTCAATTTAGGAAAAAANAAC 457418
QY 794 TTTATTTGAAGTATTTATTTGATTTCGCCCTTAAATTTCTGGAAGCTCGTGATCTCTAAAGGC 853
Db 457417 TTTTTCAGAGTTTTCATTTGATACCTTATGAAATTTGTGAATTTTCGAGATCTCTNAAAA 457358
QY 854 CTATACAGCTTGCACGTACAGGAAGATTAAGGTTTTCACTGGGAATTTGATGATCCATAC 913
Db 457357 TTATATAACAAGCCCGCACAGGCCAAATATCTGATTTTTTACTGGTATACAAATGTACATAC 457298
QY 914 GAACCAACCAATTAATGGTGAGATAGTAATTAAGATGAAAGAT 955
Db 457297 GAAACTCTATAACACCTGATGTTCTTTTAAAGGGGTACAGAT 457256

RESULT 8
US-09-949-016-1468
; Sequence 1468, Application US/09949016

Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1468
LENGTH: 1875
TYPE: DNA
ORGANISM: Human
US-09-949-016-1468

Query Match 11.4%; Score 139; DB 4; Length 1875;
Best Local Similarity 58.2%; Pred. No. 1.6e-27;
Matches 266; Conservative 0; Mismatches 185; Indels 6; Gaps 1;
QY 497 GCCTGTGTCGTATGATAACAGGACTCAGTGGTTTCAGGGAAAAGTACTCTTGCATGTGCA 556
DB 154 GTTGACACAGTTGGCTAACAGCTTGTCTGGAGCGGGAAGACTACTGTGAGCATGGCC 213
QY 557 CTGAGTCTGAGTTGCAATTCAGAGGCCACCTCAGTATGTACTTGTGATGGTGACAACTC 616
DB 214 TTGGAGGAGTACCTGCTTTTGTCTATGATGATTCATCTGATCTGATGATGATGAT 273
QY 617 AGACATGGCTAAATAGAGATTAAGCTTTAAGGACAGAGACCGTGCAGAAAATATACGA 676
DB 274 CGTCAAGGTCTCAATAAAAACTTGGCTTTAGTCTGAGACAGAGAGAGAAATGTTTCA 333
QY 677 AGAGTTGGTGAAGTGCAGAGCTTTTGTGATGCTGTTGATGATGATGATGATGATGAT 736
DB 334 CGCATCGCAGAGTTGCTTAACTGTTTGAGATGCTGCTGATGATGATGATGATGATGAT 393
QY 737 ATATCTCCATACAGGAGAGATCG-----TGATGATGCTCGTCTACTTCCACATTTCT 790
DB 394 ATATCACCTTACCTCAGGATCGCAACATGCAAGGCAAAATTCATGAGGTCAAGTTTA 453
QY 791 AACTTTATTGAAGTATTATTGATTTGGCCCTTAAATAATTTGTAAGCTCGTGATCCTAAA 850
DB 454 CGGTTTTTTGAAGTATTGTTGATGCTCTCTGATGCTTTGTAAGAGGATGATGATGATGAT 513
QY 851 GGCCTTATACAGCTTGCAGTACAGAGAAATTAAGGTTTTCACCTGGAATTCATGATGATCA 910
DB 514 GGACTCTACAAAAAGCCCGGACAGAGAAATTAAGGTTTTCACCTGGAATTCATGATGATCA 573
QY 911 TACGAACCAACCAATTAATGTCAGATAGTAATTAAGA 947
DB 574 TATGAAAGCCAGAGGCCCTCGAGTTGGTGTCTGAAAA 610

RESULT 9
US-08-879-561-4
Sequence 4, Application US/08879561
Patent No. 5817482
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Hawkins, Phillip R.
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA: US/08/879,561
APPLICATION NUMBER: US/08/879,561
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0325 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2506 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LUNGN020
CLONE: 373887
US-08-879-561-4

Query Match 11.4%; Score 139; DB 1; Length 2506;
Best Local Similarity 58.2%; Pred. No. 1.8e-27;
Matches 266; Conservative 0; Mismatches 185; Indels 6; Gaps 1;
QY 497 GCCTGTGTCGTATGATAACAGGACTCAGTGGTTTCAGGGAAAAGTACTCTTGCATGTGCA 556
DB 189 GTTGACACAGTTGGCTAACAGCTTGTCTGGAGCGGGAAGACTACTGTGAGCATGGCC 248
QY 557 CTGAGTCTGAGTTGCAATTCAGAGGCCACCTCAGTATGTACTTGTGATGGTGACAACTC 616
DB 249 TTGGAGGAGTACCTGCTTTTGTCTATGATGATTCATGCTACACTCTGGATGGTGACAAAT 308
QY 617 AGACATGGCTTAAATAGAGATTTAAGCTTTAAGGACAGAGACCGTGCAGAAAATATACGA 676
DB 309 CGTCAAGGTCTCAATAAAAACTTGGCTTTAGTCTGAGACAGAGAGAGAAATGTTTCA 368
QY 677 AGAGTTGGTGAAGTGGCAAGCTTTTGTGATGCTGTTGATGATGATGATGATGATGATGAT 736
DB 369 CGCATCGCAGAGTTGCTTAACTGTTTGATGATGCTGCTGATGATGATGATGATGATGAT 428
QY 737 ATATCTCCATACAGGAGAGATCG-----TGATGATGCTCGTCTACTTCCACATTTCT 790
DB 429 ATATCACCTTACCTCAGGATCGCAACATGCAAGGCAAAATTCATGAGGTCAAGTTTA 488
QY 791 AACTTTATTGAAGTATTATTGATTTGGCCCTTAAATAATTTGTAAGCTCGTGATCCTAAA 850
DB 489 CCGTTTTTTGAAGTATTGTTGATGCTCTCTGATGATGTTTGTGAACAGAGGATGATCAAA 548
QY 851 GGCCTTATACAGCTTGCAGTACAGGAAAGATTAAGGTTTTCACCTGGAATTCATGATGATCA 910
DB 549 GGACTCTACAAAAAGCCCGGACAGAGAAATTAAGGTTTTCACCTGGAATTCATGATGATCA 608
QY 911 TACGAACCAACCAATTAATGTCAGATAGTAATTAAGA 947
DB 609 TATGAAAGCCAGAGGCCCTCGAGTTGGTGTCTGAAAA 645

```
RESULT 10
US-09-898-165B-10
; FILE REFERENCE: 18810-81552
; CURRENT APPLICATION NUMBER: US/09898165B
; PATENT NO. 6818428
; GENERAL INFORMATION:
; APPLICANT: Daniel H. Cohn
; APPLICANT: Muhammad Faiyaz ul Haque
; APPLICANT: Lily M. King
; APPLICANT: Deborah Krakow
; TITLE OF INVENTION: 3-Phosphoadenosine-5-Phosphosulfate
; TITLE OF INVENTION: (PAPS) Synthetase Proteins and Methods for Treating
; TITLE OF INVENTION: Osteoarthritic Disorders
; FILE REFERENCE: 18810-81552
; CURRENT APPLICATION NUMBER: US/09/898,165B
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/399,212
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1851
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-898-165B-10

Query Match      10.9%; Score 132.8; DB 4; Length 1851;
Best Local Similarity 57.1%; Pred. No. 7.6e-26;
Matches 264; Conservative 0; Mismatches 192; Indels 6; Gaps 1;

QY 496 AGGCTGTGTCGATGATGAACAGGACTCAGTGGTTTCAGGGAAAAAGTACTCTTGCATGTGC 555
Db 126 AGGATGTACCGTGTGGCTAACAGGTCTCTCTGGTGTGGGAAAAACAACATAAGCTTTGC 185
QY 556 ACTGAGTCGTGAGTTGCATTGTCAGAGGCCACTCAGTGTACTTGTGTTGGTGAACACT 615
Db 186 TTGGAAGAGTACCTTGTATCTCAGCCATCCATGTTACTCCCTGGATGGGCAATGT 245
QY 616 CAGCATGGCCCTAAATAGAGATTTAAAGCTTTAAAGGCAGAGACCGTGCAGAAAAATATACG 675
Db 246 CCGTCATGGCCCTTAAAGAACTGGGATCTCTCGCGGGACCGAGAGAGATATCCG 305
QY 676 AAGAGTTGGTGAAGTGGCAAGCTTTTTCGTGATGCTGGTGTCTATGATGATGCTAGCTT 735
Db 306 CCGGATCGCGAGGTGGCAAGCTTTTTCGTGATGCTGGTGTCTATGATGATGCTAGCTT 365
QY 736 GATATCTCCATACAGGAGATCGTGATCGATCGGTGCTCT-----ACTTCCACATTC 789
Db 366 TATCTCTCTTTTGCAAGGATCGTGATCGATCGGTGCTCT-----ACTTCCACATTC 425
QY 790 TAACTTTTATGAAGTATTATTGATTTGCCCTTAAATAATTTGTGAAGCTCGTGATCTTAA 849
Db 426 CCGGTTCTTTGAGATCTTTGATAGTCGCCCTTTAAATATCTGTGAAGCCGAGAGCTAAA 485
QY 850 AGGCTATACAGCTTGACGTACAGGAAAGATTAAGGTTTCACTGGAATGATGATCC 909
Db 486 AGGACTCTACAAACGAGCCGAGGAGAGATTAAAGGTTTACAGGATCGATCTTGA 545
QY 910 ATACGAACCAACCAATTAATGTTGAGATAGTAAATTAAGATGA 951
Db 546 CTATGAGAAACCTGAAACTCCAGAGTGTGTGCTGAAGACCAA 587

RESULT 11
US-09-898-165B-2
; Sequence 2, Application US/09898165B
; Patent No. 6818428
; GENERAL INFORMATION:
; APPLICANT: Daniel H. Cohn
; APPLICANT: Muhammad Faiyaz ul Haque
; APPLICANT: Lily M. King
; APPLICANT: Deborah Krakow
; TITLE OF INVENTION: 3-Phosphoadenosine-5-Phosphosulfate
; TITLE OF INVENTION: (PAPS) Synthetase Proteins and Methods for Treating
```

```
; TITLE OF INVENTION: Osteoarthritic Disorders
; FILE REFERENCE: 18810-81552
; CURRENT APPLICATION NUMBER: US/09/898,165B
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/399,212
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-898-165B-2

Query Match      10.9%; Score 132.8; DB 4; Length 2000;
Best Local Similarity 57.1%; Pred. No. 8e-26;
Matches 264; Conservative 0; Mismatches 192; Indels 6; Gaps 1;

QY 496 AGGCTGTGTCGATGATGAACAGGACTCAGTGGTTTCAGGGAAAAAGTACTCTTGCATGTGC 555
Db 186 AGGATGTACCGTGTGGCTAACAGGTCTCTCTGGTGTGGGAAAAACAACATAAGCTTTGC 245
QY 556 ACTGAGTCGTGAGTTGCATTGTCAGAGGCCACTCAGTGTACTTGTGTTGGTGAACACT 615
Db 246 TTGGAAGAGTACCTTGTATCTCAGCCATCCCATGTTACTCCCTGGATGGGCAATGT 305
QY 616 CAGCATGGCCCTAAATAGAGATTTAAAGCTTTAAAGGCAGAGACCGTGCAGAAAAATATACG 675
Db 306 CCGTCATGGCCCTTAAAGAACTGGGATCTCTCGCGGGACCGAGAGAGATATCCG 365
QY 676 AAGAGTTGGTGAAGTGGCAAGCTTTTTCGTGATGCTGGTGTCTATGATGATGCTAGCTT 735
Db 366 CCGGATCGCGAGGTGGCAAGCTTTTTCGTGATGCTGGTGTCTATGATGATGCTAGCTT 425
QY 736 GATATCTCCATACAGGAGATCGTGATCGATCGGTGCTCT-----ACTTCCACATTC 789
Db 426 TATCTCTCTTTTGCAAGGATCGTGATCGATCGGTGCTCT-----ACTTCCACATTC 485
QY 790 TAACTTTTATGAAGTATTATTGATTTGCCCTTAAATAATTTGTGAAGCTCGTGATCTTAA 849
Db 486 CCGGTTCTTTGAGATCTTTGATAGTCGCCCTTTAAATATCTGTGAAGCCGAGAGCTAAA 545
QY 850 AGGCTATACAGCTTGACGTACAGGAAAGATTAAGGTTTCACTGGAATGATGATCC 909
Db 546 AGGACTCTACAAACGAGCCGAGGAGAGATTAAAGGTTTACAGGATCGATCTTGA 605
QY 910 ATACGAACCAACCAATTAATGTTGAGATAGTAAATTAAGATGA 951
Db 606 CTATGAGAAACCTGAAACTCCAGAGTGTGTGCTGAAGACCAA 647

RESULT 12
US-09-543-681A-2995
; Sequence 2995, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 2995
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-2995

Query Match      10.6%; Score 128.4; DB 4; Length 681;
Best Local Similarity 55.3%; Pred. No. 7e-25;
```


Search completed: September 9, 2005, 12:44:43
Job time : 256 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 9, 2005, 09:55:25 ; Search time 830 Seconds
(without alignments)
9629.315 Million cell updates/sec

Title: US-10-829-432-3
Perfect score: 1217
Sequence: 1 ggcgtcggttcattcatca.....aaaaaaaaaaaaaaaaaaaa 1217

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7351250 seqs, 3283620254 residues
Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16:	/cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17:	/cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18:	/cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19:	/cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
20:	/cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
21:	/cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
22:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
23:	/cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
24:	/cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1217	100.0	1217	19	US-10-829-432-3
2	1037.4	85.2	1291	20	US-10-425-115-116953
3	859.2	70.6	1335	18	US-10-425-114-941
4	704.8	57.9	1498	20	US-10-425-115-116950
5	604.2	49.6	1498	19	US-10-767-701-9302
6	555	45.6	1319	20	US-10-739-930-5324
7	387	31.8	548	19	US-10-437-963-72904

8	359.6	29.5	928	19	US-10-829-432-9	Sequence 9, Appli
9	352.2	28.9	915	19	US-10-437-963-51779	Sequence 51779, A
c 10	350.2	28.8	378	17	US-10-260-238-5283	Sequence 5283, Ap
11	342	28.1	1256	18	US-10-424-599-62905	Sequence 62905, A
12	338.8	27.8	936	19	US-10-829-432-7	Sequence 7, Appli
13	336	26.1	2611	20	US-10-425-115-62889	Sequence 62889, A
14	333	27.4	1180	18	US-10-425-114-18247	Sequence 18247, A
15	333	27.4	1312	18	US-10-425-114-1150	Sequence 1150, Ap
16	333	27.4	1764	20	US-10-425-115-135127	Sequence 135127,
17	330.4	27.1	981	18	US-10-424-599-119620	Sequence 119620,
18	328.2	27.0	1116	18	US-10-425-114-9652	Sequence 9652, Ap
19	328.2	27.0	1302	18	US-10-425-114-27400	Sequence 27400, A
20	316.2	26.0	890	19	US-10-829-432-1	Sequence 1, Appli
21	316.2	26.0	1028	18	US-10-425-114-16540	Sequence 16540, A
22	294.4	24.2	831	9	US-09-938-842A-710	Sequence 710, App
23	294.4	24.2	831	11	US-09-938-842A-710	Sequence 710, App
24	277	22.8	1132	18	US-10-424-599-119621	Sequence 119621,
25	264.8	21.8	687	17	US-10-260-238-869	Sequence 869, App
26	263.6	21.7	920	16	US-10-342-224-65	Sequence 65, Appl
27	245	20.1	262	9	US-09-923-876-1680	Sequence 1680, Ap
28	245	20.1	262	10	US-09-923-876-1680	Sequence 1680, Ap
29	238.6	19.6	615	20	US-10-425-115-126988	Sequence 126988,
c 30	233.4	19.2	595	17	US-10-260-238-4843	Sequence 4843, Ap
31	219.2	18.0	1592	19	US-10-437-963-45866	Sequence 45866, A
32	218.2	17.9	722	19	US-10-767-701-29593	Sequence 29593, A
33	211.4	17.4	635	20	US-10-425-115-35035	Sequence 35035, A
34	199.8	16.4	1821	20	US-10-425-115-35036	Sequence 35036, A
35	197.2	16.2	431	19	US-10-829-432-5	Sequence 5, Appli
36	188.6	15.5	328	11	US-09-732-627A-1323	Sequence 1323, Ap
37	185.4	15.2	897	18	US-10-425-114-16007	Sequence 16007, A
38	181.6	14.9	1113	20	US-10-425-115-62888	Sequence 62888, A
39	179	14.7	597	17	US-10-282-122A-35733	Sequence 35733, A
40	170.2	14.0	603	17	US-10-282-122A-15460	Sequence 15460, A
41	168.6	13.9	642	17	US-10-282-122A-41542	Sequence 41542, A
42	162.8	13.4	521	19	US-10-829-432-11	Sequence 11, Appl
43	161.4	13.3	600	17	US-10-369-493-41371	Sequence 41371, A
44	159.6	13.1	615	22	US-10-724-972A-1295	Sequence 1295, Ap
45	158.2	13.0	597	17	US-10-282-122A-34804	Sequence 34804, A

ALIGNMENTS

RESULT 1
US-10-829-432-3
; Sequence 3, Application US/10829432
; Publication No. US20040177401A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio
; APPLICANT: Allen, Stephen
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-B
; CURRENT APPLICATION NUMBER: US/10/829,432
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/720,384A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1217
; TYPE: DNA
; ORGANISM: Zea mays
US-10-829-432-3

Query Match 100.0%; Score 1217; DB 19; Length 1217;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GCGTCGCGTTTCATTCATCAACAGAACCTCTGTCTACACACGACGACACACC 60
|||||

Db 1 GCGTCGGTTTCATTTTCATCAATCAAAACAGAACCTCTGTTGTCACACACGACGCAACACCACC 60
Qy 61 GAGCCAGCCCGCCGAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 120
Db 61 GAGCCAGCCCGCCGAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 120
Qy 121 GAGCGGAGCGCTCGCGCTCATCTCGTAAATCCACAGCGCGCGCTCGCGCTCGCGCTCGCGCTCG 180
Db 121 GAGCGGAGCGCTCGCGCTCATCTCGTAAATCCACAGCGCGCGCTCGCGCTCGCGCTCGCGCTCG 180
Qy 181 CCTCACCTTAGCGATGCGGCACCTCCCGCGCTCGTGTATCGATGCGCTCATCTCCCGTTC 240
Db 181 CCTCACCTTAGCGATGCGGCACCTCCCGCGCTCGTGTATCGATGCGCTCATCTCCCGTTC 240
Qy 241 CTCACACTCTCCCGGGTCTCGCCAGTGTAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db 241 CTCACACTCTCCCGGGTCTCGCCAGTGTAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Qy 301 TCGCGGTACGCACTGCGCCAGCGCGCAATGGCGGTGGGTGGCGCGCGCGCGCGCGCGCGCGCG 360
Db 301 TCGCGGTACGCACTGCGCCAGCGCGCAATGGCGGTGGGTGGCGCGCGCGCGCGCGCGCGCG 360
Qy 361 AGCAGCGCCCGGGAGGCCCCCAGCCAGCTGAGGAGAGCGCTGTAATGTGCAACAT 420
Db 361 AGCAGCGCCCGGGAGGCCCCCAGCCAGCTGAGGAGAGCGCTGTAATGTGCAACAT 420
Qy 421 TGGGAAATCGACTAATATTTATGACCAATTTGATGACCAATTTGATGACCAATTTGATGACCA 480
Db 421 TGGGAAATCGACTAATATTTATGACCAATTTGATGACCAATTTGATGACCAATTTGATGACCA 480
Qy 481 ATTGCTGGGCAAAAAGGCTGTGCTGATGATAACAGGACTCAGTGGTTTCAGGGAAAAG 540
Db 481 ATTGCTGGGCAAAAAGGCTGTGCTGATGATAACAGGACTCAGTGGTTTCAGGGAAAAG 540
Qy 541 TACTCTTGATGTGACATGAGTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 600
Db 541 TACTCTTGATGTGACATGAGTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 600
Qy 601 TGATGTTGACAACTTCAGACATGGCTTAAATAGAGATTTAAGCTTTAAGGCAAGAGCCG 660
Db 601 TGATGTTGACAACTTCAGACATGGCTTAAATAGAGATTTAAGCTTTAAGGCAAGAGCCG 660
Qy 661 TGCAGAAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTGTGATGCTGTGATGAT 720
Db 661 TGCAGAAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTGTGATGCTGTGATGATGAT 720
Qy 721 ATGCAATGCTAGCTTGATATCTCCATACAGGAGATGATGATGATGATGATGATGATGATGAT 780
Db 721 ATGCAATGCTAGCTTGATATCTCCATACAGGAGATGATGATGATGATGATGATGATGATGAT 780
Qy 781 TCCACATTTCACTTTTATGAAAGTATTTATGATTTGCCCTTAAATAATTTGTGAAGCTCG 840
Db 781 TCCACATTTCACTTTTATGAAAGTATTTATGATTTGCCCTTAAATAATTTGTGAAGCTCG 840
Qy 841 TGATCCTAAAGGCTATACAAGCTTGCACTACAGGAAAGATTAAGGTTTCACTGGAAT 900
Db 841 TGATCCTAAAGGCTATACAAGCTTGCACTACAGGAAAGATTAAGGTTTCACTGGAAT 900
Qy 901 TGATGATCCATACGAACCAATTAATGGTGAGATAGTAAATTAAGATGAAGATGAGGA 960
Db 901 TGATGATCCATACGAACCAATTAATGGTGAGATAGTAAATTAAGATGAAGATGAGGA 960
Qy 961 ATGGCCCTTCAACCAAGCAATGGCCAGCAAGTTCATGCTACCTTGAAGAAAACGGATA 1020
Db 961 ATGGCCCTTCAACCAAGCAATGGCCAGCAAGTTCATGCTACCTTGAAGAAAACGGATA 1020
Qy 1021 TTTGCAAGCTTAGTATGATATTTTGAAGAAGATGATCTGATCTGTGTGTCATTACT 1080
Db 1021 TTTGCAAGCTTAGTATGATATTTTGAAGAAGATGATCTGATCTGTGTGTCATTACT 1080
Qy 1081 TGTGGACAAATAAGATCTGTGTTGGTGCACATGATTAAGAGGCATCAACATGTAGGAAG 1140
Db 1081 TGTGGACAAATAAGATCTGTGTTGGTGCACATGATTAAGAGGCATCAACATGTAGGAAG 1140

Qy 1141 TAACAGAGGTACGGTTTCATTTCAGAAACGGATATGGATTTCATTGTTTAAAAA 1200
Db 1141 TAACAGAGGTACGGTTTCATTTCAGAAACGGATATGGATTTCATTGTTTAAAAA 1200
Qy 1201 AAAAAA 1217
Db 1201 AAAAAA 1217
RESULT 2
US-10-425-115-116953
; Sequence 116953, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kowalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 116953
; LENGTH: 1291
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_38151C.1
US-10-425-115-116953
Query Match 85.2%; Score 1037.4; DB 20; Length 1291;
Best Local Similarity 95.8%; Pred. No. 7.6e-279;
Matches 1083; Conservative 0; Mismatches 44; Indels 3; Gaps 2;
Qy 64 CCCAGCGCCCGCCAGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCCAG 123
Db 2 CGAGCAACACCGAGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCCAGCGCCAG 61
Qy 124 GCGAGCGTCCCGTCACTCTCGTAAATCCACAGCGCGCGCTCCCGTCTCCAGGCT 183
Db 62 GCGAGCGTCCCGTCACTCTCGTAAATCCACAGCGCGCGCTCCCGTCTCCAGGCT 121
Qy 184 CACCCCTAGCATGCGCCACTCCCGCGCTCGTATCCATGCGCTCACTCCCGTCTC 243
Db 122 CACCCCTAGCATGCGCCACTCCCGCGCTCGTATCCATGCGCTCACTCCCGTCTC 181
Qy 244 ACACCTTTCCCGGGTCTTCGCCAGTGTAGTGGCGCGCGCGAGGGAGGGCCGCGGTGC 303
Db 182 ACACCTTTCCCGGGCTTCGCCAGTGTAGTGGCGCGCGCGAGGGAGGGCCGCGGTGC 241
Qy 304 CGGTACGACATGCGCACCGCGGCTTGGCGGTGGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 363
Db 242 CGGTACGACATGCGCACCGCGGCTTGGCGGTGGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 301
Qy 364 AGCGCC -CGGGAGCGCCCGCACCGCCAGTGAAGGAGAGCCTGTATGTCAACATTTG 422
Db 302 AGCGCCAGCGGAGCGCCCGCACCGCCAGTGAAGGAGAGCCTGTATGTCAACATTTG 361
Qy 423 GGAATCGACTAATATTTATGGCACAATTCCTTTGTCGCAATCTGATGACAGAAAT 482
Db 362 GGAATCGACTAATATTTATGGCACAATTCCTTTGTCGCAATCTGATGACAGAAAT 421
Qy 483 TGCTGGGCAAAAAGGCTGTGCTGATGATAACAGGACTCAGTGGTTTCAGGAAAAGTA 542
Db 422 TGCTGGGCAAAAAGGCTGTGCTGATGATAACAGGACTCAGTGGTTTCAGGAAAAGTA 481
Qy 543 CTCTTGATGTGCACTGAGTGTGATGTTGCAATTTGCAGAGGCCACCTCACGTATGTACTTG 602
Db 482 CTCTTGATGTGCACTGAGTGTGATGTTGCAATTTGCAGAGGCCACCTCACGTATGTACTTG 541

Qy	603	ATGCTGAACAACCTCAGACATGGCCCTAAATAAGAGATTTTAAGCTTTAAGCGCAGAAGACC	GTG	662
Db	542	ATGCTGAACAACCTCAGACATGGCCCTAAATAAGAGATTTTAAGCTTTAAGCGCAGAAGACC	GTG	601
Qy	663	CAGAAAATATACGAAGAGTGGTCAAGCTGGCAAAGCTTTTTTGCTGATGCTGGTGTGCATAT		722
Db	602	CAGAAAATATACGAAGAGTGGTCAAGCTGGCAAAGCTTTTTTGCTGATGCTGGTGTGCATAT		661
Qy	723	GCAATTGCTAGCTTGATATCTCCATACAGGAGAGATCGTGATGTCATGCCGTCTACTTTC		782
Db	662	GCAATTGCTAGCTTGATATCTCCATACAGGAGAGATCGTGATGTCATGCCGTCTACTTTC		721
Qy	783	CACATTCTAACCTTTATTGAAGTATTTATTGATTTGGCCCCCTAAAAATTTGTGTGAAGCTCGTG		842
Db	722	CACATTCTAACCTTTATTGAAGTATTTATTGATTTGGCCCCCTAAAAATTTGTGTGAAGCTCGTG		781
Qy	843	ATCCTAAGCGCTATACAAGCTTGCACGTACAGGAAGATTTAAAGGTTTTCACCTGGGAATTC		902
Db	782	ATCCTAAGCGCTATACAAGCTTGCACGTACAGGAAGATTTAAAGGTTTTCACCTGGGAATTC		841
Qy	903	ATGATCCATACGAAACCAACCAATTAATGGTGAGATAGTAGTAATTAAGATGAAAGATGAGGAAT		962
Db	842	ATGATCCATACGAAACCAACCAATTAATGGTGAGATAGTAGTAATTAAGATGAAAGATGAGGAAT		901
Qy	963	GCCCTTCACCCAAAGCAATGCCACGAAGTTCTATGCTACCTTTGAAGAAAAACGGATATT		1022
Db	902	GCCCTTCACCCAAAGCAATGCCACGAAGTTCTATGCTACCTTTGAAGAAAAACGGATATT		961
Qy	1023	TGCAAGCTTAGTATATGTAATTTTGAAGAGATTCGATCTGATTCTTTGTGTGCTCAATTACTTCG		1082
Db	962	TGCAAGCTTAGTATATGTAATTTTGAAGAGATTCGATCTGATTCTTTGTGTGCTCAATTACTTCG		1021
Qy	1083	TGGACACAATAAGATCTGTTGTTGGTCACATGAATAAAGGCATCAACATGTAGGAAGTA		1142
Db	1022	TGGACACAATAAGATCTGTTGTTGGTCACATGAATAAAGGCATCAACATGTAGGAAGTA		1081
Qy	1143	ACGAAGGTAACGGTTCATT--CAGAAACGGATATGGATTCAATTCGTTTAA		1190
Db	1082	ACGAAGCTTCGATTTCATCTGTCATCAACCAATTCAGGATGATTCGGTTAA		1131

RESULT 3

```

US-10-425-114-941
; Sequence 941, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 941
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700097023_FLI
US-10-425-114-941

```

RESULT 4

US-10-425-115-116950
; Sequence 116950, Application US/10425115

Query Match	49.6%;	Score 604.2;	DB 19;	Length 945;
Best Local Similarity	91.6%;	Pred. No. 7.7e-158;		
Matches 664;	Conservative 0;	Mismatches 53;	Indels 8;	Gaps 2
Qy	466	ATCTGATAGACAGAAAATTCGTGGACAAAAAGCGTGTCTGTATCGGATAACAGGACTCAG	525	
Db	1	ATCTGATAGACAGAAGTTGCTGGCACAAAAGCGTGTCTGTGTGGATAACAGGACTCAG	60	
Qy	526	TGGTTTCAGGAAAAGTACTCTTGCACTGTGCACCTCAGTTCGTGACATTGCAGAGGCCA	585	
Db	61	TGGTTTCAGGAAAAGTACTCTTGCTGTGCACCTCAGTCATGAGTTGCATTATAGAGGCCA	120	
Qy	586	CCTCACGTTATGTACTTCGATGGTGCAACCTCAGACATGGCCCTAAATAGAGATTTTAAAGCTT	645	
Db	121	CATCACGTTATGTACTTCGATGGTGCAACCTTAGACATGGCCCTCAATCGAGATTTTAAAGCTT	180	
Qy	646	TAAAGCGAAGACCGTGCAGAAAAATATACGAAGAGTTGGTGAAGTGGCAAGACTTTTTGC	705	
Db	181	TAAAGCGAAGACCGTGCAGAAAAATATACGAAGAGTTGGTGAAGTGGCAAGACTTTTTGC	240	

```
QY 706 TGATGCTGGTGCATATGCAATCTAGCTTGATATCTCCATACAGGAGAGATCGTGATGC 765
DB 241 CGATGCTGGTATCATATGCATCTAGCTTGATATCTCCATACAGGAGAGATCGTGATGC 300
QY 766 ATGCCGTGCTTACTTCCACATCTCAATCTTATTAAGTATTTATTTGATTTGCCCTTAAA 825
DB 301 ATGCCGTGCTTACTTCCAGATTTCACTTATTAAGTATTTATTTGATTTGCCCTTAAA 360
QY 826 AATTGCTGAAGCTCGTGATCTTAAAGCCTATACAGCTTGCACGTFACAGGAAAGATTAA 885
DB 361 AATTGCTGAAGCTCGTGATCTTAAAGCTGTACAGCTTGCACGACAGGAAAGATTAA 420
QY 886 AGTTTCACTGGAAATTTGATGATCCATACGAACCAACCAATTAATGTTGAGATAGTAATTA 945
DB 421 AGTTTCACTGGAAATTTGATGATCCATACGAACCGCCAGTTAATGTTGAGATAGTAATTA 480
QY 946 GATGAAGATGAGGATGCTTCCACCAAGCAATGGCCACAGCAAGTTCTATGCTACT 1005
DB 481 GATGAAGATGAGGATGCTTCCACCAAGCAATGGCCACAGCAAGTTCTATGCTACT 540
QY 1006 TGAAGAAAACGGATATTTGCAAGCTTAGTATATGATTTTGAAGATTTGATCTGATCT 1065
DB 541 TGAAGAAAACGGATATTTGCAAGCTTAGTATATGATTTTGAAGATTTGATCTGATCT 597
QY 1066 TGTGTGCTCATTACTTGTGGACACATAAAGATCTGTTGTTGTTGTCACATGAATAAAGGCA 1125
DB 598 TGTGTGCTCATTACTTGTGGACACATAAAGATCTGTTGTTGTTGTCACATGAATAAAGGCA 652
QY 1126 TCAACATGTAGGAGTAACAGAGGTACGGTTCATTTCAGAAACGGATATGGAATTCATTCG 1185
DB 653 TCAACATGTAGGAGTAACAGAGGTACGGTTCATTTCAGAAACGGATATGGAATTCATTCG 712
QY 1186 TTATA 1190
DB 713 TTICA 717

RESULT 6
US-10-739-930-5324
; Sequence 5324, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 5324
; LENGTH: 1319
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-CLUSTER7564_1
US-10-739-930-5324

Query Match 45.6%; Score 555; DB 20; Length 1319;
Best Local Similarity 84.7%; Pred. No. 5.3e-144;
Matches 636; Conservative 0; Mismatches 110; Indels 5; Gaps 1;

QY 378 CCCCGCACGCCAGTGAAGGAGAGCCGTGAATGTCGAACATTCGGAAATCGACTAATA 437
DB 155 CACTCGCGCCGACTGCGAGAGAGCCGTGCTGATGTCGAACATTCGGAAATCGACTAATA 214
QY 438 TTTTATGGCAATTTGCTTGTGATTTGGCAATCTGTATAGACAGAAATTTGCTGGACAAAAAG 497
DB 215 TTTATGGCATGACTGCCCAATTTGACAAACTGAGCGACAGAAATTTGCTGGGCAAAAAG 274
QY 498 GCTGTGCTGTATGATTAACAGACATCAGTGGTTCAGGGGAAAGTACTCTTGTGATGTGAC 557
DB 275 GATGTGTCATATGGATCAAGGACTCAGCGGTTCAGGGGAAAGTACCGTTGCTGTGAC 334
```

```
QY 558 TGAGTCGTGAGTTGCAATTCAGAGGCCACCTCACGTATATGTAATCTGATGGTGACAACTCA 617
DB 335 TGAGTCGTGGAATTTACACTACAGAGGCCACCAACGTATGCTCTTGTGATGGTGACAACTCA 394
QY 618 GACATGGCCTAAATAGAGATTTAAGCTTTAAGGCGAAGAACCGTGCAGAAAATATACGAA 677
DB 395 GACATGGCCTCAATCGAGATCTAAGCTTTCAAGGCGAAGAACCGTGCAGAAAATATACGAA 454
QY 678 GAGTTGGTGAAGTGGCAAGAGCTTTTGTGCTGATGCTGGTGTGCATATGCAATGCTAGCTTGA 737
DB 455 GAGTTGGTGAAGTGGCAAGAGCTTTTGTGCTGATGCTGGTGTGCATATGCAATGCTAGCTTGA 514
QY 738 TATCTCCATACAGGAGAGATCGTGATGATGCGCTGCTTACTTCCACATTTCTAACTTTA 797
DB 515 TATCTCCATACAGGAGAGACCGTGATGATGCCGAGCTTACTTCCAGATTTCTAGATTTA 574
QY 798 TTGAAGTATTTTATGATTTGCCCCCTAAATAATTTGTGAAGCTCGTGATCTCTAAAGGCTTAT 857
DB 575 TTGAAGTATTTTATGATTTGCCCCCTAGAAATTTATGTGAAGCTCGTGATCTCTAAAGGCTTAT 634
QY 858 ACAAGCTTGCACGTACAGGAAAGATTAAAGTTTTCACCTGGAATTCATGATCCATACGAAC 917
DB 635 ACAAGCTTGCACGTACAGGAAAGATTAAAGGTTTCAACCGGAGTTGATGATCCATACGAAT 694
QY 918 CACCAATTAATGCTGAGATAGTAAATTAAGATGAAGATGAGGAATGCCCTTCACCCAAAG 977
DB 695 CACCAATTAATGCTGAGATAGTAAATTAAGATGAGGAATGCCCTTCACCCAAAG 754
QY 978 CAATGGCCAAAGCAATCTTCTACTCTACCTTGAAGAAACGGATATTTGCAAGCTTAGTATA 1037
DB 755 CAATGGCCAAAGCAATCTTCTACTCTACCTTGAAGAAATGGATATTTGCAAGCTTAGTATA 814
QY 1038 TGTATT-----TTGAGAGATTTGATCTGATCTTCTGTTGTTGCTTACTTGTGGACACAT 1092
DB 815 CATACTCCAGATCCAGAAAGTTGATCTTATTTCTTCTGTTGCTTAACTAATCTGACACAGG 874
QY 1093 AAGATCTGTTGTTGTCACATGAATAAAGG 1123
DB 875 TACAATCCATTTGTTCCATCCGGAATAAAG 905

RESULT 7
US-10-437-963-72904
; Sequence 72904, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 72904
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_73237C.1
US-10-437-963-72904

Query Match 31.8%; Score 387; DB 19; Length 548;
Best Local Similarity 81.7%; Pred. No. 2.9e-97;
Matches 447; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
```

321	Qy	GGGCGCATTTGGCGGTGGGTGGCGCGCGCGCGCGGAAATGGAGCAGCCGCCCGGGGAGGCCCC	380
2	Db	GGGAGGCGGCCCGCGCCGGGAGGGATGGAGCAACCAACAACAGCTTTGCCGGGGACAGGT	61
381	Qy	CGCACAGCCCACTGAAGGAGGAAGCTGTAATGTGCGAACAATTGGGAAATCGACTAATATT	440
62	Db	TCCGCGGCCCAATAAAGAGAAGCCTTTGTGTCAACCATTGGGAAATCAACCAATATAC	121
441	Qy	TATGGCACAATTTGCTTGGATTGGACAATCTGATAGACAGAAATTCGTGGGACAAAAAGGCT	500
122	Db	TGTGGCACAAATTGCCCAATTGGACAACTCTGAGAGCGAGAACTTGTGGGGCAAAAAGGAT	181
501	Qy	GTGTGCTATGGATAACAGAACTCAGTGGTTCAGGGGAAAAGTACTCTTTCGATGTGCACTGA	560
182	Db	GTGTCAATTTGGATTCACAGGACTTCAGCGCTCAGGGGAAAAGCACTCTTGCACTGTGCACTGA	241
561	Qy	GTCTGTAGTTGCATTTGCAGAGGCCACCTCAGTATGTACTTTGATGGTGGCAACCTCAGAC	620
242	Db	ATCGGAGTTGCATTTGCAGCGGCCACCTTACTTATGCTTGTGATGGTGACACCTAAGGC	301
621	Qy	ATGGCCATAATAGAGATTTAAGCTTTAAGGCAGAGAAGCCGTGCAGAAAAATATACGAAGAG	680
302	Db	ATGGCCTAATCGAGATCTAAGCTTCAAGGCAGAGAAGCCGTGCAGAAAATATACGAAGAG	361
681	Qy	TTGGTGAAGTGCGCAAGCTTTTGTGTGATGCTGTGTGTCATATGCAATTCGTAGCTTGATAT	740
362	Db	TTGGAGAAGTGCGCAAGCTTATTCGCAGATGCTGGAAATCATATGCAATTCGTAGTTTGATAT	421
741	Qy	CTCCATACAGGAGAGATCGTGATGCATGCCGTGCTCTACTTCCACATTTCTAACTTTATTG	800
422	Db	CTCCATATAGGAGAGACCGTGATGCGCGCTGTGCTTCACAGAGCTCTAGATTTATTG	481
801	Qy	AAGTATTTATTGATTTGGCCCTTAAAAAATTTGTGAAGCTCGTGTATCTCTAAAGGCTTATACA	860
482	Db	AGTATTTATGATTTGGCACTAGAAAATTTGTGAAGCTCGTGATCTCTAAAGGCTTATACA	541
861	Qy	AGCTTGC	867
542	Db	AGCTTGC	548

RESULT 8

```

US-10-829-432-9
; Sequence 9, Application US/10829432
; Publication No. US20040177401A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio
; APPLICANT: Allen, Stephen
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-B
; CURRENT APPLICATION NUMBER: US/10/829,432
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/720,384A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 928
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-829-432-9

```

	Query Match	29.5%	Score 359.6	DB 19	Length 928
	Best Local Similarity	65.5%	Pred. No. 1.8e-89		
	Matches 558	Conservative 0	Mismatches 289	Indels 5	Gaps 2
Qy	367	GCCCGGGAGGCCCGCCACGCCCCAGTGAAGGAGAAGCCTGTATATGTCACATTGGGAA	426		
Db	77	GCCCGTCNAATGGATCAGCCATGGCAGGTATCGCAAGCTTGTCACCTCAACTTTGGGAA	136		

Qy	427	ATCGACTAATATTTTATATGGACAAATTCCTGTGATTGGACAACTCTGATATAGACAGAAATTGCT	486
Db	137	ATCGACAAACGTTCTTTGGCATGACATGTCCAATTAGGTCTAGTTTGGAGAGGCGAGAACTGCT	196
Qy	487	GGGACAAAAGGCTGTGTCTATATGGATAACAGGACTCAGTGGTTTCAGGGAAAAAGTACTCT	546
Db	197	AAATCAGNAGGTTGTGTGTGTGATATACAGGGTTAAGTGGTTACGGGAAAAGCACACT	256
Qy	547	TGCATGTGCACACTCAGTCTGTGAGTTGCATATGCAGAGCCACCTCACGPTATGTACTTGATGG	606
Db	257	AGCATGGCGCTTAAGTCGCGAGCTGCACTCCAGAGGTCTACTGACCTTACATTCTAGACGG	316
Qy	607	TGCACACCTCAGACATGGCTTAATAGAGATTAAAGCTTTAAGCGCAGAGACCGTGCAGA	666
Db	317	TGCAATCTTAAGCATGGGTTAAACCGAGACCTCTGTTCGAAGCAAGACCGTGTCTGA	376
Qy	667	AAATATACGAAGAGTTGGTGAAGTGGCAAAAGCTTTTTCGTGATGCTGGTCTCATATGCAT	726
Db	377	AAATATACGCAGATAGGAGAAGTAGCAAAAGCTGTTTGACAGATGCTGGTCTGATCTGCAT	436
Qy	727	TGCTAGCTTGATATCTCCATACACAGAGAGATCTGTGATGCATGCGCGTCTCTACTTCCACA	786
Db	437	TGCTAGCTTGATATCAACCTTACAGAAATGAAACGACGCTTGGCGCAAAATTACTGCACAA	496
Qy	787	TTCTCAACTTATTGAAGTATTATTGATTTGCCCCCTAAAAATTTGTGAAGCTCGTGATCC	846
Db	497	TTCTACATTCATCGAGGTGTTTTTGAAATGTCCCACTTGAAGTTTGTGAAGCTAGGGATCC	556
Qy	847	TAAAGGCTTATACAAAGCTTCACGCTACAGGAAAGATTAAAGTTTTCACATGGAATTGTATGA	906
Db	557	AAAAGGCTTGTAACAAGCTTGCCCGTCGAGGAAAAATCAAAAGGTTTACTCGAATTGATGA	616
Qy	907	TCCATACGAACCAATTAATCGTGAGATAGTAATTAAGATGAAAGATGAGGAATGCCCC	966
Db	617	TCCTTATGAAGCACCTTCTGACTCGGAGATAGTGATACAGTGCAGAAAGCTGGTGACTGCGC	676
Qy	967	TTCAACCCAAAGCAATGSCCAAGCAAGTTCTATGCTACTCTTGAAGAAAAACGATATTTGCA	1026
Db	677	CAGCCCTAAATCGATGGCTGATCAAGTTGTGTCATATCTTGAAGCAAAATGAGTTCATTACA	736
Qy	1027	AGCTTAG---TATATGPTATTTTGAGAAGATTGATCTGATTCCTTGTGTGCCATTACTTGT	1083
Db	737	GGAATAGAGACGTATGCTATGGATGAAAAACAATCTGAAATTTGGATTCGCCCAAGGGATGT	796
Qy	1084	GGACACAATAAGATCTGTGCTTGGTCACATGAATAA--AAGGCATCAACATGTAGGAAGT	1141
Db	797	GAAATATGAGGTAGTATTTATGTCTGAAAGAGGTGATGATGATGAGAACATATATATT	856
Qy	1142	AACAGAGGTGACGGTTCATTTCAGAAAAACGGATATGGATTCAATTCGTTTAAAAAATAAAAA	1201
Db	857	GACATAAAGATCGAACTCTGTACATCATTAATAAATTTGAAATGTTTTTGACGCAAAAAAA	916
Qy	1202	AAAAAAAAAAAAA	1213
Db	917	AAAAAAAAAAAAA	928

RESULT 9

```

: US-10-437-963-51779
: Sequence 51779, Application US/10437963
: Publication No. US20040123343A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Kovalic, David K.
: APPLICANT: Zhou, Yihua
: APPLICANT: Cao, Yongwei
: APPLICANT: Wu, Wei
: APPLICANT: Boukharov, Andrey A.
: APPLICANT: Barbazuk, Brad
: APPLICANT: Li, Ping
: TITLE OF INVENTION: Rice Nucleic Acid
: TITLE OF INVENTION: Plants and Uses T
: FILE REFERENCE: 38-21(53221)B

```



```
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 51779
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_5413C.1
US-10-437-963-51779

Query Match      28.9%; Score 352.2; DB 19; Length 915;
Best Local Similarity 69.6%; Pred. No. 2.1e-87;
Matches 492; Conservative 0; Mismatches 213; Indels 2; Gaps 1;

QY 329 GGGCGGTGGTGGCGGGGGGGAATGAGCAGCGCCCGGGGAGCCCGGCACAG- 387
Db 209 GGTGAGGCGCGGTGGCGGGGGGCGCCGGTGCACGGGAAGGACGACGACGATGT 268
QY 388 -CCCACTGAAGGAGAGCCCTGTAATGTCGAACATTTGGGAAATCGACTAATATTTATGGC 446
Db 269 TGCAGTTTTAAACAAGCTCATGACCTCAACTGTCGTTAAATCAACAACATCCGATGGC 328
QY 447 ACAATGCTTGAATGGACAACTCTGATAGACAGAAATGCTGGGACAAAAGGCTGTGCG 506
Db 329 ATGACTGCCCAGTAAACCAAGCTTGACAGACAAAAGTTGCTGAACCAAGAGGCTGTGTTG 388
QY 507 TATGATACAGACTCATGCTGTTGAGGAAAGTACTTTCATGTCGACTGAGTCGTG 566
Db 389 TTTGATCATCTGGCTTAAGTGTGTTGAGGAAAGACCCCTTGCATGTGCACTGAGCCGTG 448
QY 567 AGTTGATTCAGAGAGCCACCTCACGTATGACTTTGATGTCACAACTTCAGACATGCGC 626
Db 449 AGCTGCACTCAAGAGGCACTGACCTATGTTCTTGACGGCGACAACTCTCGGCAAGGCC 508
QY 627 TAAATAGAGATTTAAGCTTTAAGGAGAGAACCGTGCAGAAAATATACGAAGAGTTGTTG 686
Db 509 TGAACAAAGATCTCAGCTTCAAGCAAGGATCGTGCCGAAATATATCGCAGATTGGAG 568
QY 687 AGTGGCAAGCTTTTGTGATGCTGTTGATGATGATGATGATGATGATGATGATGATGAT 746
Db 569 AAGTGGCAAGCTTTTGTGATGCTGTTGATGATGATGATGATGATGATGATGATGATGAT 628
QY 747 ACAGAGAGATCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 806
Db 629 ATAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 688
QY 807 TTATGATTTGCCCTTAAAGTTTGTGAAGCTCGTGATCCTTAAAGGCTTATACAGCTTG 866
Db 689 TCCTGAATGTCCTCACTTGAAGTATGTAAGAAAGGATCCAAAGGCTGTGACAGCTTG 748
QY 867 CAGGTACAGGAAGATTAAGTTTCACTGGAAATGATGATGATGATGATGATGATGATGAT 926
Db 749 CTCGTGCGCGGCAAAATCAAGGCTTTACGGGAATGATGATGATGATGATGATGATGAT 808
QY 927 ATGTGTAGATAGTAAATTAAGATGAAGATGAGGAATGCCCTTCAACCAAGCAATGGCCA 986
Db 809 ATTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 868
QY 987 AGCAAGTTCTATGCTACCTTGAAGAAAACGGATATTTGCAAGCTTAG 1033
Db 869 ATCAAGTAGTGTCATATCTTGAAGCAATGGATTTCTTCAGAACTAG 915
```

RESULT 10

```
US-10-260-238-5283/c
; Sequence 5283, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
```

RESULT 11

```
US-10-424-599-62905
; Sequence 62905, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 62905
```

```
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Rieke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 5283
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Zea mays
US-10-260-238-5283
```

```
Query Match      28.8%; Score 350.2; DB 17; Length 378;
Best Local Similarity 98.7%; Pred. No. 4.5e-87;
Matches 374; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 245 CACTTTCGCGGGTCTGCGCAGTGATAGTGGCGCGCCGAGGGGAGGGCGCGGTGCG 304
Db 378 CACTTTCGCGGGTCTGCGCAGCGATAGTGGCGCGCCGAGGGGAGGGCGCGGTGCG 319
QY 305 CGTAGGCACTCCACCCCGGCGCATTTGGCGGTGGTGGCGCGCGCGCGCGGGAATGGAGCA 364
Db 318 CGTAGGCACTCCACCCCGGCGCATTTGGCGGTGGTGGCGCGCGCGCGCGGGAATGGAGCA 259
QY 365 GCGCC-CGGGAGAGCCCGCAGCACAGCCAGTGAAGAGAGAGCTGTAAATGTCGAACTGG 423
Db 258 GCGCCAGGGGATGCCCCGCGCACAGCCAGTGAAGAGAGAGCTGTAAATGTCGAACTGG 199
QY 424 GAAATCGACTAATATTTTATGGCAAAATGCTTCAATGGCAAAATCTGTATAGACAGAAAT 483
Db 198 GAAATCGACTAATATTTTATGGCAAAATGCTTCAATGGCAAAATCTGTATAGACAGAAAT 139
QY 484 GCTGGGCAAAAAGGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 543
Db 138 GCTGGGCAAAAAGGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 79
QY 544 TCTTGCATGTGCACTGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 603
Db 78 TCTTGCATGTGCACTGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 20
QY 604 TGGTGACAACCTCAGACAT 622
Db 19 TGGTGACAACCTCAGACAT 1
```

```
; LENGTH: 1256
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_27817C.1
US-10-424-599-62905

Query Match      28.1%; Score 342; DB 18; Length 1256;
Best Local Similarity 71.9%; Pred. No. 1.8e-84;
Matches 447; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

QY 410 ATGTCGAACATTGGGAATCGCAATATATTTATGGCAAAATGCTTGATGGACAATCT 469
Db 200 ATGTCAAATGTTGGCAATCGCAAAATATATATGTCATGACTGTCCAAATTCAGAAACAA 259

QY 470 GATAGACAGAAATTCCTGGGACAAAAGGCTGTGCTATGATGAATACAGGACTCAGTGGT 529
Db 260 GATAGACAGCAGCTGCTTCAGCAAAAAGGCTGTGTTATATGGCTAACTGGGCTCAGTGGT 319

QY 530 TCAGGGAAAAGTACTCTTGTCATGTGCACCTGAGTCTGAGTTGCAATTCAGAGGCCACCTC 589
Db 320 TCAGGAAAAGACACTATTGCAATGCTTTGAGTCGAAGCTTGCACTCCAAAGGAAAATCG 379

QY 590 ACGTATGACTTGATGGTGACAACTCAGACATGCGCTAAATPAGAGATTTAAGCTTTAAG 649
Db 380 TCTTACATCCTTGATGGTGACAAATATTCGGCATGCTTAAACCAAGATCTTAGTTTCAGA 439

QY 650 GCAGAGACCGTGCAGAAATATACGAAGAGTTGCTGAGTGGCAAGCTTTTTCCTGAT 709
Db 440 GCAGAGAGATCGTCTGAAACAAATTAAGAAGGATTTGAGGTTTTCATAGATGTTCCACTACATGT 499

QY 710 GCTGCTGTCATATGCAATGCTAGCTTGATATCTCCATACAGAGAGATCGTGATGCATGC 769
Db 500 GCTGCTGTTATTTGCACTCACTAGTTTATATCACCATACCAAAAGATAGAGATGCATGC 559

QY 770 COTGCTCTACTTCCACATTTCTAACTTTATTTGAAGTATTTATTTGATTTGCCCTTAAAT 829
Db 560 AGAGCACTAAATTCCAAAGGAGATTTTATGAGGTTTTCATAGATGTTCCACTACATGTG 619

QY 830 TGTGAAGCTCGTGAATCTTAAGGCTTATACAGCTTGCAAGCTTGCAAGGAAAGATTAAGGT 889
Db 620 TGTGAAGCTTAGGACCCCAAGGAGCTCTACAAGCTTTGCTCGAGCTGGAAAGATCAAAAGT 679

QY 890 TTCACTGGAATTTGATGATCCATACGAACCAACCAATTAATGTTGAGATAGTAAATTAAGATG 949
Db 680 TTCACTGGTATAGATGATCCATATGACCAACCGTGTAGTTGTGAGATAGTATTACAACAG 739

QY 950 AAAGATAGGAATGCGCTTCACCCAAAGCAATGGCCAAAGCAAGTTCTATGCTACCTTGA 1009
Db 740 AAAGGAAGTGACTGTGAAGTCTCCAGTGATATGGCTGAAGAAGTGATATCTTACTTTGGAG 799

QY 1010 GAAACCGGATATTTCCAGCTT 1031
Db 800 GAGAACGGATACCTGCGGGCTT 821
```

```
RESULT 12
US-10-829-432-7
; Sequence 7, Application US/10829432
; Publication No. US20040177401A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio
; APPLICANT: Allen, Stephen
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-B
; CURRENT APPLICATION NUMBER: US/10/829,432
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/720,384A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 14
```

```
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 936
; TYPE: DNA
; ORGANISM: Glycine max
US-10-829-432-7

Query Match      27.8%; Score 338.8; DB 19; Length 936;
Best Local Similarity 71.5%; Pred. No. 1.2e-83;
Matches 445; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 410 ATGTCGAACATTGGGAATCGCAATATATTTATGGCAAAATGCTTGATTTGGACAATCT 469
Db 313 ATGTCAAATGTTGGAACTCGCAAAACATTTATGTCATGACTGTGCTCAATTCAGAAACAA 372

QY 470 GATAGACAGAAATTCCTGGGACAAAAGGCTGTGCTATGATGAATACAGGACTCAGTGGT 529
Db 373 GATAGACAGCAGCTGCTTCAGCAACAAGGCTGTGTTATATGGCTAACTGGGCTCAGCGGA 432

QY 530 TCAGGGAAAAGTACTCTTGTCATGTGCACCTGAGTCTGAGTTGCAATTCAGAGGCCACCTC 589
Db 433 TCAGGAAAAGACACTATTGCAATGCTCTGAGTCAAAGCTTGCACTCCAAAGGAAAATCG 492

QY 590 ACGTATGACTTGATGGTGACAACTCAGACATGCGCTTAAATPAGAGATTTAAGCTTTAAG 649
Db 493 TCTTACATCCTTGATGGTGACAAATATTCGGCATGCTTAAACCAAGGATCTTAGTTTTAGA 552

QY 650 GCAGAGACCGTGCAGAAATATACGAAGAGTTGCTGAGTGGCAAGCTTTTTCCTGAT 709
Db 553 GCAGAGAGATCGTCTGAAACAAATTAAGAAGGATTTGAGGTTGGCAAACTCTTTTCAGAT 612

QY 710 GCTGCTGTCATATGCAATGCTAGCTTGATATCTCCATACAGAGAGATCGTGATGCATGC 769
Db 613 GCTGCTGTTATTTGCACTCACTAGTTTAAATACCAATACCAAAAGATAGAGATGCATGC 672

QY 770 CGTGTCTACTTCCACATTTCTAACTTTATTTGAAGTATTTATTTGATTTGCCCTTAAAT 829
Db 673 AGAGCACTACTTTCAAAGGAGATTTTATGAGGTTTTCATAGATGTTCCACTACATGTG 732

QY 830 TGTGAAGCTCGTGAATCTTAAGGCTTATACAGCTTGCAAGCTTGCAAGGAAAGATTAAGGT 889
Db 733 TGTGAAGCTTAGGACCCCAAGGAGCTCTACAAGCTTTGCTCGAGCTGGAAAGATCAAAAGT 792

QY 890 TTCACTGGAATTTGATGATCCATACGAACCAACCAATTAATGTTGAGATAGTAAATTAAGATG 949
Db 793 TTCACTGGTATAGATGATCCATATGACCAACCGTGTAGTTGTGAGATAGTATTACAACAG 852

QY 950 AAAGATAGGAATGCGCTTCACCCAAAGCAATGGCCAAAGCAAGTTCTATGCTACCTTGA 1009
Db 853 AAAGGAAGTGACTGTGAAGTCTCCAGTGATATGGCTGAAGAAGTGATATCTTACTTTGGAG 912

QY 1010 GAAACCGGATATTTCCAGCTT 1031
Db 913 GAGAACGGATACCTGCGGGCTT 934
```

```
RESULT 13
US-10-425-115-62889
; Sequence 62889, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 62889
; LENGTH: 2611
```



```
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: (1) (2611)
; LOCATION: (1) (2611)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_157357C.1
US-10-425-115-62889

Query Match      27.6%; Score 336; DB 20; Length 2611;
Best Local Similarity 63.8%; Pred. No. 1.3e-82;
Matches 528; Conservative 0; Mismatches 295; Indels 5; Gaps 1;

QY 206 CCGGCGCTGTCATCATGGCTCACTCCCGCTTCTCACACTCTTCGCGGGTCTCGCC 265
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1344 CCGGAACGGAGCCCGCGCGAATCTAGGTTGCTCGCTCGCCCTCTCCCGAG 1403
QY 266 AGTGATAGTGGCGCGCGAGGGGAGGCCCGCTGCGCTACGACTGCCACCGCGCC 325
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1404 ACTCCTAGCGGAGGAATTTAGGACGAGGTAAACGCGACAGAGCTGGAAGGAGCGCC 1463
QY 326 ATTGGCGGTGGTGGCGCGCGCGGCGGGAATGAGAGAGCCCGGGAGGCCCGCGAC 385
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1464 CCTCGTGAATCGCGCGGCGACCGATCCGTGGAGGAGCAGC-----CGGAGCAGCAGGA 1518
QY 386 AGCCGAGTGAAGGAGAGCGCTGTAATGTCGAACATTGGGMAATCGACTAATATTTATGG 445
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1519 GTTGAAGGCGAAGCTTTGAAATGTCATCCACTGTGCCGAAGTCAATATCTCTGG 1578
QY 446 CACAAATGCTTGAITGGACAATCTGATAGACAGAAATTTCTGGGACAAAAAGGCTGTGC 505
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1579 CATGATTGCTAGTTGGCAAGACTGATCGCAGAGACTACTCAACAGAGGCTGTGT 1638
QY 506 GTATGGATAACAGGACTCAGTGGTTCAGGGAAGTACTCTTTGCAATGTCACATGAGTGT 565
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1639 GTCTGGATCACAGGCTTAGTGGCTCAGGTAAAGTACCTTGGCATGTACATTTAGGCGGG 1698
QY 566 GAGTTGCATTGACAGAGCCACCTACGATGATGTTGATGTTGATGTTGATGTTGATGTTG 625
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1699 GAGCTCCACAGAGAGGAGCTTCGATATGTTCTTGACGGTGAATACTTAAGACATGGT 1758
QY 626 TTAATAGAGATTTAAGCTTTAAGGACAGAAAGCCGTGCAGAAAAATATACGAAGAGTTGGT 685
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1759 CTGAACAGGATCTTGGCTTTAAAGCTGAGACCGTGTCTGAAAAATATAGGAGAGTTGGT 1818
QY 686 GAAATGGCAAGCTTTTGTCTGATGCTGGTGTATATGATGCTTACATTTGATATCTCCA 745
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1819 GAAATGACAAAGTTATTTGCAGATCGGCGCTTGTATGTTATGCAAGTTTGTATATCTCCA 1878
QY 746 TACAGGAGATCGTGATGATGCGCTGCTCTACTTCCACATTTCTAACTTTATTTGAAGTA 805
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1879 TATAGGAGAGACCGTAATCTTGGCGTGCATTTGTTCTGTAGTAGTCTTCAATGAAGTT 1938
QY 806 TTTATTGATTGCCCCCTAAAAATTTGTGAAGCTCGTGATCTCTAAAGGCTTATACAAAGCTT 865
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1939 TTTCTTGAATGCTCTTGGATTTATGAGACAGAGATCCAAAGGCTCTTATAGCTT 1998
QY 866 GCACGTACAGGAAGATTAAAGGTTTCACTGGAATTTGATGATCCATACGAACCACTAAT 925
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1999 GCTCGTGAGGAAAAATAAAGGTTTACAGGAATAGATGACCTTTATGAAGCACCACCTG 2058
QY 926 AATGTTGATAGTAAATTAAGATGAAGATGAGGAATGCCCTTACCCAAAGCAATGCC 985
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2059 AATTTGATGATTTGATCAAGAGAGTAGATGTTGTTATGCCCTCCGCTGCTGAGATGCA 2118
QY 986 AAGCAAGTTCTATGCTTACCTTCAAGAAACGATATTTGCAAGCTTAG 1033
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2119 GGGCAAGTTGTACTTACCTTGAGGAAAGGCTTCTCGACGAGTAG 2166
```

RESULT 14

US-10-425-114-18247

```
; Sequence 18247, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kowalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425.114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 18247
; LENGTH: 1180
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3079-005-C3_FLI
US-10-425-114-18247
```

```
Query Match      27.4%; Score 333; DB 18; Length 1180;
Best Local Similarity 70.2%; Pred. No. 5.6e-82;
Matches 447; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 424 GAAATCGACTAATATTTTATGGCAATTTGCTTGTGATGGACAATCTGATAGACAAATTT 483
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
328 GAAGTCATCAATATCTTCTGGCATGATTTGCTAGTTGGCAAGACTGATGCCAAGCT 387
QY 484 GCTGGGACAAAAAGCTGTGTGATGATTAACAGGACTCAGTGGTTTCAGGAAAAAGTAC 543
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
388 ACTCAACAGAAAGCTGTGTGTCTGGATCACAGGCTTAGTGGCTCAGGTAAAAGTAC 447
QY 544 TCTTGCATGTGCACCTGAGTGTGAGTTGCAATTCAGAGGCCACCTCACGTATGTATTGA 603
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
448 CTTGCGATGTACATTAGSCGGGAGCTCCACACAAGAGGGAAGCTTGATATATTTCTTGA 507
QY 604 TGGTGACACCTCAGACATGSCCTTAATAGAGATTTAAGCTTTAAGGACAGACCGTGC 663
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
508 CGGTGATAACTTAAGACATGGTCTGAACAAGACTTTGGGCTTTAAGCTGAAGATCGTGC 567
QY 664 AGAAAAATATACGAAGAGTTGGTGAAGTGGCAAAAGCTTTTCTGCTGATGCTGTGTCATATG 723
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
568 TGAATAATATAAGGAGAGTTGGTGAAGTAGCAAAAGTTATTTTCAGACGCTGGCCTGTATG 627
QY 724 CATTGCTAGCTTGATATCTCCATACAGGAGAGATCGTGATGCATGCCGCTGCTCTACTTCC 783
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
628 TATTGCAAGTTTGATATCTCCATATAGGAGACCGTGAATCTTTCGCGTGCACCTGTGTC 687
QY 784 ACATTCTAACTTTATGAAATTTATTTGATTTGCCCTTAAAAATTTGTGAAGCTCGTGA 843
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
688 GGATGGTAGCTTTTATGAAAGTTTCTTGAATATGTCTCTTGGAAATTTATGTGAAGCAAGGA 747
QY 844 TCCTAAAGGCTTATACAGCTTGCACTGACAGGAAAGATTAAAGCTTTTCACTGGAAATGA 903
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
748 CCCAAGGCTCTTATAGCTTGTCTGCTGGGAAAAATAAAGGTTTTTACAGGAATAGA 807
QY 904 TGATCCATACCAACCACTAATTAATGTTGATAGTATTAATTAAGATGAAGATGAGGAATG 963
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
808 TGACCTTATGAAGCACCACCTGAATTTGAGATTTGAGATCAAGGAAATAGATGGTGTATG 867
QY 964 CCCTTCAACCAAGCAATGCGCAAGCAAGTTCTATGCTACCTTGAAGAAAAACGGATATTT 1023
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
868 CCCTTCGCTTCTGACATGGCGGCAAGTGGTTACTTTACCTTGAAGGAAAGGCTTCTCT 927
QY 1024 GCAAGCTTAGTATATGTTTGTAGAGATTTGATCTG 1060
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
928 GCACGAGTAGGCTGACCGCTGAATCGAAGCTACATCCG 964
```

RESULT 15

Query Match	27.4%	Score 333	DB 18	Length 1312
Best Local Similarity	63.4%	Pred. No. 5.9e-82		
Matches 528	Conservative 0	Mismatches 300	Indels 5	Gaps 1
Qy	206	CCGCGCGTCGTGATCCATGGCTCACTCCCGCTTCTCACACTCTTCGCGGGTCTCGCC	265	
Db	212	CCGGAACCGAGCCCGCGCGAATCGAGGGTTGCTTCGGCTCGCCCCCTCTCCCGAG	271	
Qy	266	AGTGATAGTGGCGCGCGAGGGGAGGCGCGTGGCTAGCGCACTGCGCACCGCGGC	325	
Db	272	ACTCTAGCGCAGGAATTAGGACGAGGGTAACAGGCGACAGAGGCTCGGAAGGACGCC	331	
Qy	326	ATTGGCGGTGGTGGCGCGCGCGGCGGAATGGAGCAGCGCGCGGGAGCGCCCGCAC	385	
Db	332	CTTCGTCAAAATGCGCGCGACCGATCCGTGGAGGAGCAGCGGAGCACGCGAGGAGTTA	391	
Qy	386	AGCCACGTGAAGAGAGACCTGTAA-----TGTGCAACATTTGGGAAATCGACTAATATT	440	
Db	392	GTCTAGTTGAAGGGCAAGCTTTGAAAATGTCCATCCACTGTCCGGAAGTCATCAAAATATCT	451	
Qy	441	TATGGCACAAATGTGCTGATTGGACAACTCTGATAGACAGAAATGTCTGGGACAAAAGGCT	500	
Db	452	TCGGCATGATTGCTAGTTGGCAAGCTGATCGCCAGAAGCTACTCAACCAAGAAGGCT	511	
Qy	501	GTGTCGTATGGATAACAGGACTCAGTGGTTCAGGAAAGTACTCTTGCATGTGCACATGA	560	
Db	512	GTGTTGCTGGATCACAGGCTTAGTGGCTCAGGTAAAGTACCTTGGCATGTACATTAG	571	
Qy	561	GTGCTGAGTTGCATTGCAGAGGCCACCTCAGTATGTACTTGTATGTTGTCACAACTCAGAC	620	
Db	572	GCGGGAGCTCCACACAGAGGGAGCTTGCATATGTTCTTTCGCGTGATAACTTAAAGAC	631	
Qy	621	ATGCGCTAAATAGAGATTTAAGCTTTTAAGGCAGAGACCGTGCAGAAAATATACCAAGAG	680	
Db	632	ATGCTCTGAACAGGATCTTGCGCTTTAAAGCTGAAGACCGTGTGAAATATTAAGGAGAG	691	
Qy	681	TTGTTGAAGTGGCAAGACTTTTTCGTGATGCTGGTGTGCATATGCATTCCTAGCTTGATAT	740	
Db	692	TTGTTGAAGTAGCAAGTTATTTTCAGAGCGGGCCCTTGTATGTTATTCGAAGTTTGATAT	751	
Qy	741	CTCCATACAGGAGATCGTGATGCATGCCGTGCTCTACTTCCCAATTCTTAACCTTTATG	800	
Db	752	CTCCATATAGGAGAGACCGTGAATCTTCCCGCTGCATTGTTATCCGATGGTAGCTTTATTG	811	
Qy	801	AAGTATTTATGATTGCCCCATAAATTTCTGAGCTCGTGCATCTTAAGGCCCTATACA	860	
Db	812	AAGTTTTCTTGAAATATGTTGCTTTGAAATTTATGTGAAGCAAGGAGACCCAAAGGCCCTCTATA	871	
Qy	861	AGCTTGCACGTACAGGAAGATTTAAGGTTTTCACTGGAAATTTGATGATCCATACGAACCAAC	920	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 9, 2005, 09:31:45 ; Search time 3856 Seconds
(without alignments)
12013.541 Million cell updates/sec

Title: US-10-829-432-3
Perfect score: 1217
Sequence: 1 gcgtccgttctattcatca.....aaaaaaaaaaaaaaaaaaaa 1217

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	701.6	57.6	720	7	COS28976 3530 1.19
2	590	48.5	675	2	BE357876 DG1_22_A0
3	583.4	47.9	691	2	BE355111 DG1_113_B
4	495.8	40.7	1075	7	CK205911 FGAS01747
5	486.6	40.0	617	6	CD863007 AZO1.105F
6	477.2	39.2	645	7	COS32501 3530 1.21
7	468	38.5	614	6	CA204452 SCAGFL108
8	455.4	37.4	799	7	CNI25017 RHOH1_8_H
9	446.8	36.7	809	6	CA766480 AF53-RpF
10	442.2	36.3	860	7	CR285854 CR285854
11	422.2	34.7	746	2	BE361884 DG1_82_H0
12	415	34.1	673	6	CA182186 SCBST310
13	411.2	33.8	525	6	CA502007 WHE4041_D
14	409.2	33.6	721	2	AW922946 DG1_47_E0
15	400.2	32.9	699	2	BE361874 DG1_82_G0
16	395.4	32.5	1071	7	CK212359 FGAS02423
17	389	32.0	673	6	CD924390 G750.112M
18	381	31.3	496	6	CA497171 WHE3223_H
19	371.6	30.5	866	9	CG159926 PUIGF42TB
20	371.6	30.5	922	9	CG454522 PU1X88TD
21	369	30.3	609	9	CG801367 1118020G0
22	356.2	29.3	671	7	CF633032 zmrws48_0
23	355.4	29.2	708	6	CA762202 BR060004B
24	352.2	28.9	915	9	CL981208 OsIFCC035

C	25	347	28.5	728	6	CD864184	CD864184 AZO1.109F
	26	344.8	28.3	421	2	BF484142	BF484142 WHE1788_G
	27	336.4	27.6	867	7	CO095086	CO095086 GR_Eal7T
	28	334.4	27.5	487	6	CA720307	CA720307 wkmzn.pk0
C	29	333	27.4	629	7	COS32500	COS32500 3530_1_21
	30	332.2	27.3	493	6	CD864183	CD864183 AZO1.109F
C	31	330.6	27.2	627	2	AW052991	AW052991 614077E08
C	32	323.6	26.6	720	7	CK827430	CK827430 zmrsub1_0
C	33	318.8	26.2	591	4	BG873962	BG873962 MEST44-F0
C	34	315.8	25.9	774	4	BM817061	BM817061 HC01C03_T
C	35	311.8	25.6	641	6	CD927759	CD927759 GR45.103B
C	36	311.2	25.6	754	6	CA921172	CA921172 EST638890
	37	310.2	25.5	797	7	CNI29765	CNI29765 RHOH1_37
	38	309.8	25.5	803	7	CO476247	CO476247 QSO0611_B
C	39	309.2	25.4	902	7	CV245034	CV245034 WS0256_B2
	40	309	25.4	696	6	CA085472	CA085472 SCJLAM209
	41	308.2	25.3	837	7	CO200697	CO200697 R1CNI2_1
	42	307.6	25.3	970	3	AY109629	AY109629 Zea mays
C	43	307	25.2	957	7	CV260804	CV260804 WS02015_B
	44	304.6	25.0	639	7	COS29312	COS29312 3530_1_19
	45	304.4	25.0	536	2	AW560397	AW560397 EST315445

ALIGNMENTS

RESULT 1
COS28976 720 bp mRNA linear EST 15-JUL-2004
LOCUS 3530_1_191_1_D07.Y_1 3530 - Full length cDNA library created by
DEFINITION Invitrogen from multiple tissues Zea mays cDNA, mRNA sequence.
ACCESSION COS28976
VERSION COS28976.1 GI:50333850
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 720)
Walbot, V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 3530_1_191_1 row: D column: 07.
Location/Qualifiers
1. 720
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="B73"
/db_xref="taxon:4577"
/tissue_type="multiple"
/dev_stage="varies by tissue"
/lab_host="DH10B"
/clone_lib="3530 - Full length cDNA library created by
Invitrogen from multiple tissues"
/note="Organ: silks, husks, ears, pollen, shoot tips,
leaf, root tips, whole seed, embryo; Vector: pCMV-SPORT
6.1; Site 1: EcorV; Site 2: NotI; Maize Gene Discovery
Project contracted with Invitrogen to produce a
normalized, full length library in a pSPORT vector. This
is a Gateway compatible vector, permitting clone movement
to new vector backbones for expression in diverse host
cells using recombination rather than restriction enzymes.
Details of the vector and sequencing primers are available
at ZmDB in the EST library description tables. poly(A) +


```

|||||
301 ATGCCGTGCTCTACTTCCAGATTCTAACTTATTGAAGTATTATTGATTTGCCCTTAA 360
|||||
826 AATTGTGTGAAGCTCGTGATCTTAAAGGCTTATACAGCTTGCACGTACAGGAAGATTAA 885
|||||
361 AATTGTGTGAAGCTCGTGATCTTAAAGGCTTGTACAGCTTGCACGACAGGAAGATTAA 420
|||||
886 AGGTTTCACTGGAATGTAGTATCCATACCAACCAATTAATGTTGAGATAGTAATTAA 945
|||||
421 AGGTTTCACTGGAATGTAGTATCCATACCAACCAATTAATGTTGAGATAGTAATTAC 480
|||||
946 GATGAAAGATGAGGAATGCCCTTACCCAAAGCAATGCGCAAGCTTCTAGTACT 1005
|||||
481 GATGAAAGATGCGGAATGCCCTTACCCAAAGCAATGCGCAAGCTTCTATCATCT 540
|||||
1006 TCAAGAAACCGATATTTGCAAGCTTAGTATGTTTGAAGATGATGATCT 1065
|||||
541 TGAAGAAACCGATATTTGCAAGCTTAGTATGTTTGAAGATGATGATCT 597
|||||
1066 TGTGTGTCCTACTTACCTTGTGGACACAATAGATCTGTTGTTGTCATGAATAAAAGGCA 1125
|||||
598 TGTGTGTCCTACTTACCTTGTGGACACAATAGATCTGTTGTTGTCATGAATAAAAGGCA 651
|||||
1126 TCAACATGTAGGAAGTAACAGAG 1149
|||||
652 TCAACACATAGCAAGTAACAGAGG 675

```

```

RESULT 3
BE355111
LOCUS
DEFINITION
DGI_113 B07.g1_A002 Dark Grown 1 (DGI) Sorghum bicolor cDNA, mRNA
sequence.
ACCESSION
BE355111
VERSION
BE355111.1 GI:9296311
SOURCE
Sorghum bicolor (sorghum)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACMAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 691)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and
Pratt,L.H.
An EST database from Sorghum: dark-grown seedlings
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 593 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: PolyTmix
High quality sequence start: 23
High quality sequence stop: 682
POLYA=No.

```

```

FEATURES
source
1..691
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone_lib="Dark Grown 1 (DGI)"
/note="Organ: 5-day-old dark-grown seedlings; Vector:
Lambda Zap; Site 1: XhoI; Site 2: EcoRI; The library was
made from poly-A RNA in the cloning vector lambda ZAP II.
Clones to be sequenced were prepared by mass excision."

```

ORIGIN

Query Match 47.9%; Score 583.4; DB 2; Length 691;

```

Best Local Similarity 92.8%; Pred. No. 1.3e-135;
Matches 636; Conservative 0; Mismatches 41; Indels 8; Gaps 2;
QY 473 AGACAGAAATTTGCTGGGACAAAGGCTGTGTCGTATGATTAACAGGACTCAGTGTTC 532
DB 1 AGACAGAAATTTGCTGGGACAAAGGCTGTGTCGTATGATTAACAGGACTCAGTGTTC 60
QY 533 GGGAAAGCTACTCTTTGTCATGTCAGTCTGTCAGTGTGTCATGTCAGAGGCCACCTCAGC 592
DB 61 GGGAAAGCTACTCTTTGTCATGTCAGTCTGTCAGTGTGTCATGTCAGAGGCCACCTCAGC 120
QY 593 TATGTACTTGTATGTCACAACTCAGACATGCGCTTAAATAGAGATTTTAAGCTTTAAGGCA 652
DB 121 TATGTACTTGTATGTCACAACTCAGACATGCGCTTAAATAGAGATTTTAAGCTTTAAGGCA 180
QY 653 GAAGACCGTCGAGAAATATACGAAGAGTTGGTGAAGTGGCAAAAGCTTTTTCCTATGCT 712
DB 181 GAAGACCGTCGAGAAATATACGAAGAGTTGGTGAAGTGGCAAAAGCTTTTTCCTATGCT 240
QY 713 GGTGTATATGATTTGCTAGCTTGTATATCTCATACAGAGAGATCGTGATGCGCCGT 772
DB 241 GGTATCATATGATTTGCTAGCTTGTATATCTCATACAGAGAGATCGTGATGCGCCGT 300
QY 773 GCTCTACTTCCAGATTTCTAACTTTATTGAAGTATTTATTGATTTTCCCTTAAAAATTTGT 832
DB 301 GCTCTACTTCCAGATTTCTAACTTTATTGAAGTATTTATTGATTTTCCCTTAAAAATTTGT 360
QY 833 GAAGCTCGTATCTTAAAGGCTTATACAAAGCTTGCACGTACAGGAAAGATTAAAGGTTTC 892
DB 361 GAAGCTCGTATCTTAAAGGCTTGTACAAAGCTTGCACGTACAGGAAAGATTAAAGGTTTC 420
QY 893 ACTGGAATGATGATTCATACCAAGCAATTAATGTTGATAGATAGTAATTAAGATGAAA 952
DB 421 ACTGGAATGATGATTCATACCAAGCAATTAATGTTGATAGATAGTAATTAAGATGAAA 480
QY 953 GATGAGGAATGCCCTTACCCAAAGCAATGCCAAGCAAGTTCTATGTCATCTTGAAGAA 1012
DB 481 GATGAGGAATGCCCTTACCCAAAGCAATGCCAAGCAAGTTCTATGTCATCTTGAAGAA 540
QY 1013 AACGGATATTTGCAAGCTTAGTATATGTTTGAAGAGATGATCTGATCTTGTGTGT 1072
DB 541 AACGGATATTTGCAAGCTTAGTATATGTTTGAAGAGATGATCTGATCTTGTGTGT 597
QY 1073 CCATTACTTGTGGACACAATAAGATCTGTTGTCATGATGAATAAAGGCATCAACAT 1132
DB 598 CCATTACTTGTGGACACAATAAGATCTGTTGTCATGATGAATAAAGGCATCAACAT 652
QY 1133 GTAGGAAGTAACAGAAGGTACGGTT 1157
DB 653 ATAGCAAGTAACAGAGGCGCAGTTT 677

```

```

RESULT 4
CK205911/c
LOCUS
DEFINITION
FGAS017472 Triticum aestivum FGAS: Library 5 GATE 7 Triticum
aestivum cDNA, mRNA sequence.
ACCESSION
CK205911
VERSION
CK205911.1 GI:39568301
KEYWORDS
EST.
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 1075)
REFERENCE
1 Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,
Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroché,A.,
Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilsson,D.,
Penniket,C., Roach,J.L. and Sarhan,F.
Functional Genomics of Abiotic Stress in Wheat and Canola Crops
Unpublished (2003)
Contact: Wm L Crosby

```

Bioinformatics
University of Saskatchewan, Department of Computer Science
1C101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada
Tel: 306 966 1769
Fax: 306 966 2033
Email: fgas_est@cs.usask.ca
This sequence is the direct result of the Base calling software
Phred (default parameters). It is the raw base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [9,754].
Plate: L5B004 row: J column: 15.

FEATURES

source

1..1075
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"
/note="Vector: pCMV.SPORT6; Crown and developmental stages
of spike formation in wheat cultivar Norstar. 4 mRNA
populations were combined before constructing the library.
The first mRNA population is from 1cm crown sections after
30 days of cold acclimation. The second is from 1cm crown
sections after 11 days of deacclimation (before
deacclimation plants were fully vernalized for 49 days).
The third is from different developmental stages of spike
formation (5 to 50mm) that still have not emerged from the
leaf (dissection required). The last is from different
developmental stages of spike and seed formation after
having emerged from the leaf (visible). First strand
synthesis in this library was done in the presence of
methylated dCTP thereby protecting from internal cleavage
with NotI."

ORIGIN

Query Match 40.7%; Score 495.8; DB 7; Length 1075;
Best Local Similarity 79.9%; Pred. No. 1.5e-113;
Matches 635; Conservative 0; Mismatches 152; Indels 8; Gaps 4;
QY 334 GTGGTGGCGCGCGCGGAATGGAGCAGCGCGCGGGAGGCGCGCAGCAGCCAGT 393
DB 1021 GGGGGTCCCGGGAACCGGATCGAACACCCCTCAGGACGGGGCCCTCCGAAGCCCA 962
QY 394 GAAGGAGAGCCTGTAATGTGCAACATGGGAAATCGACTAATATTTATGGCAATG 453
DB 961 AGTGGAAAAGCCTTAAATCCACATGGG-AATCCATTAATTTATGGCATGCTG 903
QY 454 CTTGATTGGCAATCTGATAGACAGAAATTCCTGGGACAAAAGCCTGTGCTATGGAT 513
DB 902 CCAATTTGGACAACTGAGCGACAGAAATTCCTGGGGCCAAAGGATGTTTCATTTGGAT 843
QY 514 AACAGCACTCAGTGTTCAGGAAAAGTACTCTTCGATGTGCATGTGACGTGAGTTGCA 573
DB 842 CACAGACTCAGCGTTTCAGGAAAAGTAC-CTTCCTGTGCACTGA-TCGGGGATTACA 785
QY 574 TTGCAGAGCCACCTCACGTATGTACTTGTATGTCGACACCTCAGACATGGCTAAATAG 633
DB 784 CTACAGAGGCCACCAACAGTATGCTTGNATGGTGACACCTCAGACATGGCTCAATCG 725
QY 634 AGATTTAAGCTTTAAGGCAAGAGCCGTGCAGAAAATATACAAAGATTTGGTGAATGGC 693
DB 724 AGATCTAAGCTTCAAGGCAAGAGCCGTAGAGAAAATATACAAAGATTTGGTGAATGGC 665
QY 694 AAAGCTTTTGTGATGCTGGTGTATATGATTCGATGATCTCCATACAGAG 753
DB 664 AAAGCTTTTGTGATGCTGGTACCATATGCAATGCTAGTTGATATCTCCATACAGAG 605
QY 754 AGATCGTGATGATGCGCGTCTACTTCCATCTTAACCTTTATTTAGATTTATTTGA 813
DB 604 AGACCGTGATGATGCGGAGCTTACTTCCAGATTTCTAGATTTATTTAGATTTATGA 545
QY 814 TTTGCCCTTAAAAATTTTGAAGCTCGTGATCCTTAAAGCCCTATACAAAGCTTGACAGTAC 873

DB 544 TTTGCCACTAGAAATTATGTGAAGCTCGTGTATCTCTAAAGGGTTATACAAGCTTGACGCAC 485
QY 874 AGGAAGATTAAAGTTTCACTGGAATGTATGATCCATACCAACCAATTAATGTTGA 933
DB 484 AGGAAGATTAAAGGGTTACCGGAGTTGATGATCCATACCAAGTGAATAGTGA 425
QY 934 GATAGTAATTAAGATGAAAGATGAGGAATGCCCTTCAACCCCAAGCAATGGCCAAAGCAAGT 993
DB 424 GATAGTAATTAAGATGGAAGTGGGAATGCCCTTCAACCAAGCAATGGCCCAAGCAAGT 365
QY 994 TCTATGCTACCTTCAAGAAACCGGATATTTTCAAGCTTAGTATATGTAATTTTGAAG 1048
DB 364 TCTATCTACCTTCGAGAAGATGATATTTTCAGGCTTAGCACATACCTCCAGATCCAG 305
QY 1049 AAGATTGATCGATTCGTGTGCTCCATTACTTCTGGACACAAATGAATCTGTTGTTGTT 1108
DB 304 AAGATTGATCTTATTTCTGTGTCCATACTCATGGACACAGGTACCAATCAATTTGGTTC 245
QY 1109 CACATGAATAAAGG 1123
DB 244 CATCCGGATTAAG 230
CD863007 617 bp mRNA linear EST 11-JUL-2003
AZOL105F21F010130 AZOL1 Triticum aestivum cDNA clone AZOL105F21,
mRNA sequence.
CD863007
CD863007.1 GI:32546823
EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 617)
AUTHORS Genoplatte, a major partnership french program in plant genomics
TITLE Genoplatte.
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplatte
Genoplatte
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplatte' (<http://www.genoplatte.com>
and <http://genoplatte-info.infobiogen.fr>).
FEATURES
source
1..617
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone_lib="AZOL105F21"
/tissue_type="leaf"
/clone_lib="AZOL1"
ORIGIN
Query Match 40.0%; Score 486.6; DB 6; Length 617;
Best Local Similarity 88.0%; Pred. No. 2.9e-111;
Matches 543; Conservative 0; Mismatches 69; Indels 5; Gaps 1;
QY 476 CAGAAATTCCTGGGACAAAAGGCTGTCTGATGATAACAGGACTCAGTGGTTTCAGGG 535
DB 1 CAGAAATTCCTGGGACAAAAGGATGTCTATATGATCAGAGACTCAGCGGTTTCAGGG 60
QY 536 AAAAGTACTCTTGCATGTGCATGAGTCGTGAGTTGCAATTCAGAGGCCCACTCAAGTAT 595
DB 61 AAAAGTACCGTTGCTGTGCACTGAGTCGGGAATTTACACTACAGAGGCCCACTCAAGTAT 120
QY 596 GTACTTGTGATGATGACCACTCAGACATGGCCCTAAATAGAGATTTAAGCTTTAAGCAGNA 655

```

121 GTTCTTGGTGGTACAACTCAGACATGGCTTAATCGAGATCTAAGCTTCAAGGCAGAA 180
656 GACCGTGCGAGAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTTCGTCTGCTGCT 715
181 GACCGTGACGAAGAAATATACGAAGAGTTGGAGAGTGGCAAGCTTTTTCAGATGCTGCT 240
716 GTCATATGCAATGCTAGCTTGATATCTCCATACAGGAGAGATCGTGATGCGCGTCT 775
241 ACCATATGCAATGCTAGCTTGATATCTCCATACAGGAGAGCGTGATGCGCGAGCT 300
776 CTACTTCCACATCTTAACCTTTATGAAGTATTTATGATTTGCTGCGCTTAAAGTTGTGAA 835
301 CTACTTCCAGATCTAGATTTATGAAGTATTTATGATTTGCTGCGCTTAAAGTTGTGAA 360
836 GCTCGTGATCCTTAAGGCTTATACAGCTTGACGCTACAGGAAAGATTAAGGTTTCACT 895
361 GCTCGTGATCCTTAAGGCTTATACAGCTTGACGCTACAGGAAAGATTAAGGTTTCACT 420
896 GGAATTTGATGATCCATACCAACCAATTAATGTTGATAGTAGTAATTAAGATGAAGAT 955
421 GCGTTGATGATCCATACCAATCACCAGTGATAGTAGTAGTAATTAAGATGAAGAT 480
956 GAGGAATGCTTCCACCCAAAGCAATGCGCAAGCAAGTTCTATGCTTACCTTGAAGAAAC 1015
481 GCGGAATGCTTCCACCCAAAGCAATGCGCAAGCAAGTTCTATGCTTACCTTGAAGAAAC 540
1016 GGATATTTCCAGCTTAGTATATGT-----ATTTTGAGAGATTTGATCTGTTGCT 1070
541 GGATATTTCCAGCTTAGTATATATCTCCATATCAGAAAGATTTGAATTTCTTCTGT 600
1071 GTCCATTACTTGTGGAC 1087
601 GTCCATACCTTATGGAC 617
```

```

RESULT 6
LOCUS      C0532501
DEFINITION 3530_1_213_1_H05.Y.1_3530 - Full length cDNA library created by
            Invitrogen from multiple tissues Zea mays cDNA, mRNA sequence.
ACCESSION  C0532501
VERSION    C0532501.1 GI:50337375
KEYWORDS   EST.
SOURCE     Zea mays
ORGANISM   Zea mays

REFERENCE  1 (bases 1 to 645)
AUTHORS   Walbot V.
TITLE      Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL    Unpublished (1999)
COMMENT    Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 3530.1.213.1 row: H column: 05.

FEATURES   source
            1..645
            /organism="Zea mays"
            /mol_type="mRNA"
            /cultivar="B73"
            /db_xref="taxon:4577"
            /tissue_type="multiple"
            /dev_stage="varies by tissue"
            /lab_host="DH10B"
            /clone_lib="3530 - Full length cDNA library created by
            Invitrogen from multiple tissues"
```

/note="Organ: silks, husks, ears, pollen, shoot tips, leaf, root tips, whole seed, embryo; Vector: pCMV-SPORT 6.1; Site1: EcoRV; Site 2: NotI; Maize Gene Discovery Project Contracted with Invitrogen to produce a normalized, full length library in a pSPORT vector. This is a Gateway compatible vector, permitting clone movement to new vector backbones for expression in diverse host cells using recombination rather than restriction enzymes. Details of the vector and sequencing primers are available at ZmDB in the EST library description tables. poly(A)+ mRNA was prepared by Invitrogen, and equimolar amounts of RNA from each of the 12 tissue samples were mixed together for selection of mRNA with a 5' cap. After synthesis of cDNA, a normalization step was conducted against the mixture of RNA sources. This step effected a 20X to 80X reduction in common transcript types. Tissues prepared: 1. just emerging silks; 2. inner husks from ears of sample #1; 3. 20 dap aleurone; 4. immature tassels, stages from 1-2 mm to 1-2 cm; 5. 2 mm to 2 cm ears; 6. pollen; 7. 1 cm vegetative shoot tips from 15 day old seedlings; all leaves with an expanded or partially expanded sheath were removed; 8. mature leaf tissue; 9. 0.5 cm long root tips from 15 day old seedlings; 10. 10 dap whole seed; 11. 12 dap endosperm and embryo; 12. 17 dap endosperm and embryo. All of the sequenced clones in project 3530 will be archived at the University of Arizona along with the Unigene clones from the Maize Gene Discovery EST sequencing projects. Clones can be ordered through the ZmDB web site or directly from the University of Arizona (<http://www.genome.arizona.edu/orders/>). High density filters containing over 18,000 clones can also be ordered from the University of Arizona."

ORIGIN

```

Query Match      39.2%; Score 477.2; DB 7; Length 645;
Best Local Similarity 81.6%; Pred. No. 6.7e-109;
Matches 641; Conservative 0; Mismatches 3; Indels 142; Gaps 2;

QY      64  CCCAGCGCCCGCCAGCCAGCGGCCAACGCGCAACGCCCTCTCCAGCCGAC 123
DB      1  CCCAGCGCCCGCCAGCCAGCGGCCAACGCGCAACGCCCTCTCCAGCCGAC 60

QY      124 GCCGACGCTCGCGCTCATCTCTGTAATCCACAGCGCGCGCTCCCGCTCTCCAGGCT 183
DB      61 GCCGACGCTCGCGCTCATCTCTGTAATCCACAGCGCGCGCTCCCGCTCTCCAGGCT 120

QY      184 CACCCCTAGCATGCGCCACTCCCGCGCTCGTGATCCATGCGCTCACTCCCGTTCTCTC 243
DB      121 CACCCCTAGCATGCGCCACTCCCGCGCTCGTGATCCATGCGCTCACTCCCGTTCTCTC 180

QY      244 ACACCTTCCCGGGTCTCGCCAGTATAGTGGGCGCGCGAGGGGAGGCGCGGTGC 303
DB      181 ACACCTTCCCGGGTCTCGCCAGTATAGTGGGCGCGCGAGGGGAGGCGCGGTGC 240

QY      304 CGGTACGACATGCGCCACCGCGCATTTGGCGGTGGGTGGCGCGCGCGGCGGAAATGGAGC 363
DB      241 CGGTACGACATGCGCCACCGCGCATTTGGCGGTGGGTGGCGCGCGCGGCGGAAATGGAGC 300

QY      364 AGCGCC-CGGGAGGCGCGCCGACACGCCAGTGAAGAGAGAGCGCTGTATATGCGAACAATTG 422
DB      301 AGCGCCAGCGGGATGCGCGCACAGCC----- 327

QY      423 GGAATCGACTAATATTTTATGGCACATTTGCTTGTGGACAACTCTGATAGACAGAAAT 482
DB      328 ----- 327

QY      483 TGCTGGGAGAAAAGGCTGTGCTATGGATAACAGGACTCAGTGGTTTCAGGGAAGTA 542
DB      328 -----CAGGGAAGTA 339

QY      543 CTCCTGCAATGCTAGTGTGATGTCATATGAGAGGCGACCTACGATGATGTTGTTG 602
DB      340 CTCCTGCAATGCTAGTGTGATGTCATATGAGAGGCGACCTACGATGATGTTGTTG 399
```


QY 603 ATGGTGACAACTCAGACATGCGCTAAATAGAGATTTAAGCTTTTAAGCGAGAACCGTG 662
 |||||
 Db 400 ATGGTGACAACTCAGACATGCGCTAAATAGAGATTTAAGCTTTTAAGCGAGAACCGTG 459
 |||||
 QY 663 CAGAAAATATACGAAGAGTTGGTGAAGTGGCAAAAGCTTTTCTGCTGATGCTGGTGTCATAT 722
 |||||
 Db 460 CAGAAAATATACGAAGAGTTGGTGAAGTGGCAAAAGCTTTTCTGCTGATGCTGGTGTCATAT 519
 |||||
 QY 723 GCATTGCTAGCTGTATATCTCATACAGGAGATGCTGATGCTGCTGCTTACTTC 782
 |||||
 Db 520 GCATTGCTAGCTGTATATCTCATACAGGAGATGCTGATGCTGCTGCTTACTTC 579
 |||||
 QY 783 CACATTCTAACTTTATTGAAGTATTATTATTGCTGCTTAAATTTTGTGAAGCTCGTG 842
 |||||
 Db 580 CACATTCTAACTTTATTGAAGTATTATTATTGCTGCTTAAATTTTGTGAAGCTCGTG 639
 |||||
 QY 843 ATCCTTA 848
 |||||
 Db 640 ATCCTTA 645

RESULT 7
 CA204452
 LOCUS SCAGFL1086H06.9 FL1 Saccharum officinarum cDNA clone SCAGFL1086H06
 DEFINITION 5' mRNA sequence.
 ACCESSION CA204452
 VERSION CA204452.1 GI:35241732
 KEYWORDS EST.
 SOURCE Saccharum officinarum

ORGANISM
 Saccharum officinarum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
 complex.

REFERENCE
 1 (bases 1 to 614)
 Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
 The libraries that made SUCEST
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
 Contact: Arruda P
 Centro de Biologia Molecular e Engenharia Genetica
 Universidade Estadual de Campinas
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil
 Tel: 55 19 3788 1137
 Fax: 55 19 3788 1089
 Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bcccenter.fcav.unesp.br
 Plate: 086 row: H column: 06
 Seq primer: T7 Promoter Primer.
 Location/Qualifiers
 1..614
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCAGFL1086H06"
 /lab_host="DH10B"
 /clone_lib="FL1"
 /note="Organ: Inflorescence at beginning of development
 (1cm-long); Vector: pSport1; Site 1: SalI; Site 2: NotI;
 An unidirectional cDNA library generated from
 [Inflorescence at beginning of development (1cm-long)].
 cDNA was prepared from polyA+ mRNA using SuperScript
 Plasmid System Kit (Invitrogen). The double-strand cDNAs
 were fractionated in a sepharose CL-2B 40cm-columns and
 fragments sizing between 0.8 and 1.5 Kb were
 directionally cloned into the vector. Details
 of each
 source of RNA and library construction can be obtained at
 http://sucest.lad.ic.unicamp.br/public"

FEATURES

source
 CNI25017 799 bp mRNA linear EST 01-APR-2004
 LOCUS RH0H1_8_H03.bl_A002 Acid- and alkaline-treated roots Sorghum
 DEFINITION bicolor cDNA clone RH0H1_8_H03_A002 3', mRNA sequence.
 ACCESSION CNI25017
 VERSION CNI25017.1 GI:45949036
 KEYWORDS EST.
 SOURCE Sorghum bicolor (sorghum)
 ORGANISM Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 799)
 Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S., Klein,R.R., Liang,C.,
 Sun,F., Sullivan,R., Lim,S., Eastman,A. and Pratt,L.H.
 An EST database from Sorghum: acid- and alkaline-treated roots
 Unpublished (2003)
 Contact: Cordonnier-Pratt MM
 Other ESTs: RH0H1_8_H03.g1_A002
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860

ORIGIN

Fax: 706 583 0210
 Email: mmpratt@uga.edu
 Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.
 Seq primer: Sug3-14 (TAGCTAGCGCGCGGACCC)
 POLI=Yes.

FEATURES

Location/Qualifiers
 1..799
 /organism="Sorghum bicolor"
 /mol_type="mRNA"
 /cultivar="BTx623"
 /db_xref="taxon:4558"
 /clone="RH01.8 H03 A002"
 /lab_host="DH10B-T1 phage-resistant E. coli"
 /clone_lib="Acid- and alkaline-treated roots"
 /note="Organ: Root; Vector: pME18S-FL3; Site.1: XhoI; Site.2: XhoI; The library was prepared from polyA+ RNA from 8-day-old roots harvested from BTx623 sorghum seedlings grown in hydroponic culture. HCl was added to a pH of 3.0 to some seedlings, KOH to a pH of 9.0 for others. Roots were harvested 3, 12 and 27 hr after addition and pooled for RNA preparation. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG, 3-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."

ORIGIN

Query Match 37.4%; Score 455.4; DB 7; Length 799;
 Best Local Similarity 90.4%; Pred. No. 2.1e-103;
 Matches 511; Conservative 0; Mismatches 46; Indels 8; Gaps 2;

QY 626 CTAAATAGAGATTAAAGCTTTTAAAGCAGAGACCGTGCAGAAATATACGAGAGTTGGT 685
 Db 1 CTCATCGAGATTAAAGCTTTTAAAGCAGAGACCGTGCAGAAATATACGAGAGTTGGT 60

QY 686 GAAGTGGCAAGCTTTTCTGATGCTGGTGTCTATATGATTCATGCTAGCTTGATATCTCCA 745
 Db 61 GAAGTGGCAAGCTTTTCTGATGCTGGTGTCTATATGATTCATGCTAGCTTGATATCTCCA 120

QY 746 TACAGAGAGATCGTGATGATCGCTGCTCTACTTCCACATTCATTAATGGAATGA 805
 Db 121 TACAGAGAGATCGTGATGATCGCTGCTCTACTTCCAGATTCATTAATGGAATGA 180

QY 806 TTTATTGATTTCCCTTAAATTTTGAAGCTCGTGATCCTTAAAGGCTATACAGCTT 865
 Db 181 TTTATGATTTCCCTTAAATTTTGAAGCTCGTGATCCTTAAAGGCTATACAGCTT 240

QY 866 GCACGTACAGGAAGATTAAAGGTTTCACTGGAATTTGATGATCCATACGAACCCCAATT 925
 Db 241 GCACGCACAGGAAGATTAAAGGTTTCACTGGAATTTGATGATCCATACGAACCCCAATT 300

QY 926 AATGTGAGATAGTAATTAAGATGAAGATGAGGAATGCCCTTACCCCAAGCAATGGCC 985
 Db 301 AATGTGAGATAGTAATTAAGATGAAGATGAGGAATGCCCTTACCCCAAGCAATGGCC 360

QY 986 AAGCAAGTTCTATACCTTGAAGAGAACTGATATTTGCAAGCTTAGTATGATTTT 1045
 Db 361 AAGCAAGTTCTATACCTTGAAGAGAACTGATATTTGCAAGCTTAGTATGATTTT 417

QY 1046 GAGAAGATTGATCTGATTTCTGTGTCCATTTGTTGGACACAAATAGATCTGTGTT 1105
 Db 418 GAGAATATTGATCTGATTTCTGTGTCCATTTGTTGGACACAAATAGATCTGTGTT 475

QY 1106 GGTCAATGAATAAAGGCATCAACATGTGAGAGGTAAAGAGGTACGTTTCATTCAGA 1165
 Db 476 --GCCATGAATAAAGGCATCAACATGTGAGAGGTAAAGAGGTACGTTTCATTCAGA 532

QY 1166 AACGATATGATTCATTTCGTTAA 1190
 Db 533 ACGGAATACAGATTCATTTCGTTCA 557

RESULT 9
 CA766480
 LOCUS
 DEFINITION
 CAY766480 809 bp mRNA linear EST 08-JAN-2003
 AF53-Rpf.11.K01.T7.013.ab1.IRRI Drought Stress Panicle Library
 Oryza sativa (indica cultivar-group) cDNA clone C0004081 5' similar to Adenylisulfate kinase, chloroplast precursor (BC 2.7.1.1.25) (APS kinase) (Adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-phosphotransferase), mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 CAY766480.2 GI:27548315
 EST.
 Oryza sativa (indica cultivar-group)
 Oryza sativa (indica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 809)
 REFERENCE
 AUTHORS
 Bennet, J., Arumugam, K., Lafitte, R., Wen, J., Rudd, S. and Bruskewich, R.M.
 TITLE
 JOURNAL
 COMMENT
 IRRI Drought Stress Panicle cDNA Library
 Unpublished (2002)
 On Dec 2, 2002 this sequence version replaced gi:25995735.
 Contact: Richard Bruskewich
 Biometrics and Bioinformatics Unit
 International Rice Research Institute
 DAPO 7777, Metro Manila, Philippines
 Tel: +63-2-845-0563
 Fax: +63-2-845-0606
 Email: r.bruskewich@cgiar.org
 International Rice Information System (IRIS);
 http://www.iris.irri.org; D0204080
 Assignment of putative function to the sequence by S. Rudd of the Munich Information Center for Protein Sequences
 (http://mips.gsf.de)
 Plate: 11 row: K column: 01.
 Location/Qualifiers
 1..809
 /organism="Oryza sativa (indica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="IR64"
 /db_xref="taxon:39946"
 /clone="C0004081"
 /tissue_type="Panicles"
 /dev_stage="Flowering"
 /clone_lib="IRRI Drought Stress Panicle Library"
 /note="Vector: pBluescript II SK+; Water stress was applied by not watering for 4 consecutive days. Panicles were collected from control (well watered) and stressed plants at 2 days before heading, at heading, 50% flowering and 4 days after 50% flowering."

FEATURES
 source
 location/Qualifiers
 1..809
 /organism="Oryza sativa (indica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="IR64"
 /db_xref="taxon:39946"
 /clone="C0004081"
 /tissue_type="Panicles"
 /dev_stage="Flowering"
 /clone_lib="IRRI Drought Stress Panicle Library"
 /note="Vector: pBluescript II SK+; Water stress was applied by not watering for 4 consecutive days. Panicles were collected from control (well watered) and stressed plants at 2 days before heading, at heading, 50% flowering and 4 days after 50% flowering."

ORIGIN
 Query Match 36.7%; Score 446.8; DB 6; Length 809;
 Best Local Similarity 86.5%; Pred. No. 3.1e-101;
 Matches 517; Conservative 0; Mismatches 77; Indels 4; Gaps 2;

QY 532 AGGGAAGTACTCTTGCATGCTGCTGAGTTCGATTCAGAGCCACCTAC 591
 Db 2 AGGGAAGTACTCTTGCATGCTGCTGAGTTCGATTCAGAGCCACCTAC 61

QY 592 GTATGTACTTGTGTGACCACTCAGACATGGCCCTAAATAGATTTAAGCTTTAAGGC 651
 Db 62 TTAATGCTTGTGTGACCACTCAGACATGGCCCTAAATAGATTTAAGCTTTAAGGC 121

QY 652 AGAAGACCGTCAGAAAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTTCGTGATGC 711
 Db 122 AGAAGACCGTCAGAAAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTTCGTGATGC 181

High quality sequence start: 30
High quality sequence stop: 719
POLYA=No.

FEATURES

source

Location/Qualifiers

1..746
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone_lib="Dark Grown 1 (DGI)"
/note="Organ: 5-day-old dark-grown seedlings; Vector:
lambda Zap; Site: 1: XhoI; Site 2: EcoRI; The library was
made from poly-A RNA in the cloning vector lambda Zap II.
Clones to be sequenced were prepared by mass excision."

ORIGIN

Query Match 34.7%; Score 422.2; DB 2; Length 746;
Best Local Similarity 90.3%; Pred. No. 4.8e-95;
Matches 476; Conservative 0; Mismatches 43; Indels 8; Gaps 2;
QY 664 AGAAATATACGAAGATTGGTGAAGTGGCAAGCTTTTGTGATGCTGGTGTCTATATG 723
DB 1 AGAAATATACGAAGATTGGTGAAGTGGCAAGCTTTTGTGATGCTGGTGTCTATATG 60
QY 724 CATTCGTAGTATATCCATACAGAGAGATCGTGATGATGCGGTGCTCTACTTCC 783
DB 61 CATTCGTAGTATATCCATACAGAGAGATCGTGATGATGCGGTGCTCTACTTCC 120
QY 784 ACATTTCTAATTTATTGAAGTATTATGATTTGCCCTTAAATTTGTGAGCTCGTA 843
DB 121 AGATTTCTAATTTATTGAAGTATTATGATTTGCCCTTAAATTTGTGAGCTCGTA 180
QY 844 TCCTAAAGCCCTATACAAAGCTTGACAGTACAGAAAGATTAAAGGTTTCACTGGAATTGA 903
DB 181 TCCTAAAGCTTGATACAGCTTGACAGTACAGAAAGATTAAAGGTTTCACTGGAATTGA 240
QY 904 TGATCCATACGAACCAACCAATTAATGGTGAGATAGTAATTAAGATGAAGATGAGGAATG 963
DB 241 TGATCCATACGAACCCAGCTTAATGGTGAGATAGTAATTAAGATGAAGATGAGGAATG 300
QY 964 CCCTTCAACCAAGCAATGGCAAGCTTCTATGCTACCTTGAAGAAACGGATATTT 1023
DB 301 CCCTTCAACCAAGCAATGGCAAGCTTCTATGCTACCTTGAAGAAACGGATATTT 360
QY 1024 GCAAGCTTAGTATATGATTTTGAAGAAGATTGATCTGATCTTGTGTGTCATTAATTGT 1083
DB 361 GCAAGCTTAGCAT---ATTGAGATATTGATCTGATCTTGTGTGTCATTAATTGT 417
QY 1084 GGACACAATAAGATCTGTTGTTGTCATGATGAATAAAGGCATCAACATGTAGGAAGTAA 1143
DB 418 GGACACAATAATGATCTGTTG-----GCCATGAATAAAGGCATCAACATAGCAAGTAA 472
QY 1144 CAGAGGTACGGTTTCAATTCAGAAACGGATATGATTCATTCGTTTAA 1190
DB 473 CAGAGGCGCAGTTTGTTCAGAAACGGATATGATTCATTCGTTTCA 519

RESULT 12
CA182186

LOCUS SCBGST3108C02.g ST3 Saccharum officinarum cDNA clone SCBGST3108C02
DEFINITION 5', mRNA sequence.

ACCESSION CA182186.1 GI:35117704

VERSION Saccharum officinarum

KEYWORDS EST.

SOURCE Saccharum officinarum

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.

REFERENCE 1 (bases 1 to 673)

AUTHORS Vettore, A.L., da Silva, P.R., Kemper, E.L. and Arruda, P.

TITLE The libraries that made SUCEST

JOURNAL
COMMENT

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccc.org.br
Plate: 108 row: C column: 02
Seq primer: T7 Promoter Primer.

FEATURES

source

1..673
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCBGST3108C02"
/lab_host="DH10B"
/clone_lib="ST3"
/note="Organ: Fourth apical stalk internodes of adult
plants; Vector: pSPori1; Site 1: SalI; Site 2: NotI; An
unidirectional cDNA library generated from [Fourth apical
stalk internodes of adult plants]. cDNA was prepared from
polyA+ mRNA using SuperScript plasmid System Kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Query Match 34.1%; Score 415; DB 6; Length 673;
Best Local Similarity 79.6%; Pred. No. 3.1e-93;
Matches 571; Conservative 0; Mismatches 100; Indels 46; Gaps 5;
QY 73 CGGCCAGCAGCAGCGGCGCAACCGCAAGGCAACACCTCTCTCAGCCCGACGCCACGCT 132
DB 2 CGGCCAGCAGCAGCGGCGCAACCGCTTTCGGTCCGACGCTGACGCGGGAGCGGCCACCG 61
QY 133 CGCCGCTATCTCTGTAATCCACAGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 191
DB 62 CGCGGTGGCGCCCGTAATCATAGCGCGCGGTCTCTCTCTCTCTCTCTCTCTCTCTCT 121
QY 192 GCGATGCGCCACTCCCGCGCTCGTGATCCATGCGCTCACTCCCGTTCCTTCACTCTT 251
DB 122 ACGATGCGCGCTCTGAGCTCGTGATCCATGCGCTCGCCAGCGGTGCTGGCGCC--- 178
QY 252 CCGCGGGTCTCGCCAGTGATGTCGGCGCGCGGAGGGGAGGGCGCGGTGCGGTGCGC 311
DB 179 -----GCCAGGGGGAGGAGAGGGGCGCGGTGCGGTGCGC 217
QY 312 ACTGCCACCGCGCATTTGGCGGTGGGTGCGCGCGCGCGCGGAAATCGAGCAGCGCCCG 371
DB 218 ACTGCCACCGCGCATTTGGCGGT-----GCCCGGGGATCGAGCAGCGCCAC 265
QY 372 GGGAGGCGCCCGCAGCGCCAGTGAAGGAGAGAGCTGTAAATGTGGAACATTTGGGAATCGA 431
DB 266 GG-----GGAGGGCGCGAGTGAAGGAGAGAGCTGTAAATGTGGAACATTTGGGAATCAA 317
QY 432 CTAATATTTATGCAATTTGCTTGTATGACAAATCTGATAGACAGAAATGCTGGGAC 491
DB 318 CTAATATTTATGCAATTTGCTTGTATGACAAATCTGATCGACAGAAATGCTGGGAC 377
QY 492 AAAAGGCTGTCTGTATGGAATAACAGACTCAGTGGTTTCAAGGAAAGTACTTCTTGCAT 551
DB 378 AAAAGGCT-TGTGTGTGGAATAACAGACTCAGTGGTTTCAAGGAAAGTACTTCTTGCAT 436
QY 552 GTGCACTGAGTCTGAGTTGCAATTCAGAGGCCACCTCAGCTATGTACTTGTATGATGACA 611
DB 437 GTGCACTGAGTCTGAGTTGCAATTCAGAGGCCACCTCAGCTATGTACTTGTATGATGACA 496

612	ACCTCAGACATGCCCTAAATAGAGATTAAAGCTTTTAAAGGCAGAAAGA	CCGTGCGAGAAATA	671
497	ACCTTAGACATGCCCTAAATCAGATTAAAGTTTAAAGGCAGAAAGAC	CCGTGCAGAAATA	556
672	TACCAAGAGTTGGTGAAGTGGCAAGCTTTTTGCTGATGCTGGTGTCATATG	CAATTCGCTA	731
557	TACGACGAGTTGGTGAAGTGGCAAGCTTTTTGCGCATGCTGGTATCATATG	CAATTCGCTA	616
732	GCTTGATATCTCCATACAGGAGAGATCGTGATGATGCGCGTGCTCTAC	TTCACATT	788
617	GCTTGATATCTCCATACAGGAGAGATCGTGATGATGCGCGGCTCTACT	TTCCAAAT	673

RESULT 13
CA502007
LOCUS
DEFINITION
CA502007 525 bp mRNA linear EST 14-NOV-2002 WHE4041.D11.G21.ZT wheat meiotic anther cDNA library Triticum aestivum_cDNA clone WHE4041.D11.G21.mRNA sequence.

VERSION	CA502007.1	GI:24952967
KEYWORDS	EST.	
SOURCE	Triticum aestivum (bread wheat)	
ORGANISM	Triticum aestivum	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.	

REFERENCE
1 (bases 1 to 525)
AUTHORS Anderson, O.D., Chao, S., Crossman, C., Langridge, P., Lazo, G.R.,
Pham J., Rausch, C.J., Sutton, T., Woo, J. and Wilson, C.
TITLE The structure and function of the expressed portion of the wheat
genomes - Meiotic anther cDNA library
JOURNAL Unpublished (2002)

CONTACT: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818

Email: andersnpw.usda.gov
Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: T7 primer.
Tax: J300000000

```

seq primer: 17 primer
Location/Qualifiers
1. 525
  /organism="Triticum aestivum"
  /mol_type="mRNA"
  /cultiivar="Chinese Spring"
  /db_xref="taxon:4565"
  /clone="WHE4041 D11_G21"
  /tissue_type="Anther"
  /dev_stage="Meiotic stages pre-meiosis-metaphase I"
  /lab_host="E. coli DH10B"
  /clone_lib="Wheat meiotic anther cDNA library"
  /note="vector: pSPORT1; Site 1: SalI; Site 2: NotI; Plants
were grown in a glasshouse. Anther meiotic stage was
determined by removing anthers from individual primary
florets. One anther was sacrificed for microscopic
staging, and if determined to be between (and including)
meiotic stages pre-meiosis and metaphase I, the remaining
two anthers were collected and pooled for library
construction. The tissue, total RNA, and poly(A) RNA were
prepared, cDNA synthesised, and directionally ligated into
pSPORT1 by Tim Sutton in the P Langridge Lab at the
Department of Plant Science, University of Adelaide, Waite
Campus, Australia. Average insert size 1.5Kb. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."

```

ORIGIN

```
Query Match      33.8%; Score 411.2; DB 6; Length 525;
Best Local Similarity 88.5%; Pred. No. 2.6e-92;
Matches 446; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
```

540	GTACTCTTGATGTGCACTGAGTCGTGAGTTGCAATTCGACAGAGCGCACTCACGTATGTAC	599
22	GAACCGCTGCTGTGCACTGAGTCGGGAATTACACTACAGAGGCCACACAGTATGTCC	81
600	TTGATGGTGACAACCTCAGACATGGGCTAAATAGAGATTTTAAGCTTTTAAGCGCAGAAGACC	659
82	TTGATGGTGACAACCTCAGACATGGGCTCAATCGAGACTTAAGCTTCAAGCGAGAAGACC	141
660	GTGCGAAGAAATATACGAAGAGTTGGTGAAGTGCAGAAAGCTTTTTCGTGATGCTGGTGTCA	719
142	GTACAGAAAAATATACGAAGAGTTGGTGAAGTGCAGAAAGCTTTTTCGAGATGCTGGTACCA	201
720	TATGCATTGCTAGCTTGATATCTCCATACAGGAGAGATCGTGATGATGCATGCCGTGCTCTAC	779
202	TATGCATTGCTAGTTGATATCTCCATACAGGAGAGATCGTGATGATGCATGCCAGCTCTAC	261
780	TTCCACATTTCTAATCTTATTTGAAGTATTTATGATTTGCCCTTAAATAAATTTGTGGAAGCTC	839
262	TTCCAGATTTCTAGATTTATTTGAAGTATTTATGATTTTGCCACTAGATAATTTATGTGAAGCTC	321
840	GTGATCTCTAAGGCCTATACAAAGCTTGCAAGTACAGGAAGATTTAAAGGTTTTCACATGGAA	899
322	GTGATCTCTAAGGGTTATACAAAGCTTGCAAGTACAGGAAGATTTAAAGGTTTTCACCGEAG	381
900	TTGATGATCATACGAACCAATTAATGGTGAGATAGTAAATTAAGATGAAGATGAGG	959
382	TTGATGATCATACGAATCACCAGTGAATAGTGAGATAGTAAATTAACATGGAAGGTGGGG	441
960	AATGCCCTTTCACCAAGCAATGCCAAGCAAGTCTTATGCTACCTTTGAAGAAAAACGGAT	1019
442	AATGCCCTTCCCAGGAAGCAATGGCCAGCAAGTTCTATCCTAACCTTTGAGAAATGGAAT	501
1020	ATTTCGAAGCTTAGTATATGTATT	1043
502	ATTTCGAGGCTTAGCACATATCT	525

RESULT 14

AW922946	AW922946	721 bp	mRNA	linear	EST 19-JUL-2000
LOCUS	DGI_47_E04_g1_A002	Dark Grown 1	(DGI)	Sorghum bicolor	CDNA, mRNA
DEFINITION	sequence.				
ACCESSION	AW922946				
VERSION	AW922946.1				
KEYWORDS	EST.				
SOURCE	Sorghum bicolor (sorghum)				
ORGANISM	Sorghum bicolor				

REFERENCE

AUTHORS Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and Pratt, L.H.
TITLE An EST database from Sorghum: dark-grown seedlings
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu
 Sequences have been trimmed to exclude PolyA, vector and regions
 below Phred quality 16. The threshold for highest quality sequence
 is 20.
 Seq primer: PolyTWix
 High quality sequence start: 23
 High quality sequence stop: 703
 POLYA=No.

/mol_type="mrna"
/db_xref="taxon:4558"
/clone_lib="Dark Grown 1 (DG1)"
/note="Organ: 5-day-old dark-grown seedlings; Vector:
Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was
made from poly-A RNA in the cloning vector lambda ZAP II.
Clones to be sequenced were prepared by mass excision."

ORIGIN

```
Query Match      33.6%; Score 409.2; DB 2; Length 721;
Best Local Similarity 90.1%; Pred. No. 8.9e-92;
Matches 463; Conservative 0; Mismatches 43; Indels 8; Gaps 2;

QY 677 AGAGTTGGTGAAGTGCAGAAAGCTTTTGTGATGCTGTGTGCATATGCAATGCTAGCTTG 736
Db 1 AGAGTTGGTGAAGTGCAGAAAGCTTTTGTGCGATGCTGTGTATCATATGCAATGCTAGCTTG 60

QY 737 ATATCTCCATACAGGAGATGCTGATGATGATGCGGTCTCTACTCCCAATCTCACTTT 796
Db 61 ATATCTCCATACAGGAGATGCTGATGATGATGCGGTCTCTACTCCCAATCTCACTTT 120

QY 797 ATTGAAGTATTTATGATTTGCCCTTAAAGTTTGTGAAGCTGTGATCCTAAAGCCTA 856
Db 121 ATTGAAGTATTTATGATTTGCCCTTAAAGTTTGTGAAGCTGTGATCCTAAAGCCTG 180

QY 857 TACAAGCTTGCACGTACAGGAAGATTAAAGTTTCACTGGAATTGATGATCCTACAGAA 916
Db 181 TACAAGCTTGCACGTACAGGAAGATTAAAGTTTCACTGGAATTGATGATCCTACAGAA 240

QY 917 CCACCAATTAATGGTGAGATAGTAATTAAGATGAAGATGAGGAATGCCCTTACCCCAA 976
Db 241 CGCCAGTTAATGGTGAGATAGTAATTAAGATGAAGATGAGGAATGCCCTTACCCCAA 300

QY 977 GCAATGGCCCAAGCAAGTCTATGCTACCTTGAAGAAAGCGATATTCGACCTTAGTAT 1036
Db 301 GCAATGGCCCAAGCAAGTCTATGCTACCTTGAAGAAAGCGATATTCGACCTTAGTAT 360

QY 1037 ATGTATTTTGAAGAATGATCTGATCTTGTGTGTCCATTACTTGTGGACACAATAAGA 1096
Db 361 AT---ATTGAGAAATGATCTGATCTTGTGTGTCCATTACTTGTGGACACAATAAGA 417

QY 1097 TCTGTGTGGTGCATGATGAATAAAGGCATCAACATGTAGGAAGTAACAGAGGTACGGT 1156
Db 418 TCTGTG-----GCCATGAATAAAGGCATCAACATGTAGGAAGTAACAGAGGTACGGT 472

QY 1157 TCATTGAGAAACGGATATGATTCATTCTGTTAA 1190
Db 473 TCTTGAAGAAACGGATATGATTCATTCTGTTTCA 506
```

RESULT 15
BE361874

LOCUS DG1_82_G05_g1_A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA
DEFINITION sequence.

ACCESSION BE361874

VERSION BE361874.1 GI:9303431

KEYWORDS EST.

SOURCE Sorghum bicolor (sorghum)

ORGANISM Sorghum bicolor

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE 1 (bases 1 to 699)

AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and

Pratt,L.H.

TITLE An EST database from Sorghum: dark-grown seedlings

JOURNAL Unpublished (2000)

COMMENT Contact: Cordonnier-Pratt MM

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210
Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.

Seq primer: PolyTMix

High quality sequence start: 64

High quality sequence stop: 697

POLYA=No.

FEATURES

source

Location/Qualifiers

1..699

/organism="Sorghum bicolor"

/mol_type="mrna"

/db_xref="taxon:4558"

/clone_lib="Dark Grown 1 (DG1)"

/note="Organ: 5-day-old dark-grown seedlings; Vector:

Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was

made from poly-A RNA in the cloning vector lambda ZAP II.

Clones to be sequenced were prepared by mass excision."

ORIGIN

```
Query Match      32.9%; Score 400.2; DB 2; Length 699;
Best Local Similarity 89.9%; Pred. No. 1.6e-89;
Matches 454; Conservative 0; Mismatches 43; Indels 8; Gaps 2;

QY 686 GAAGTGGCAAGCTTTTGTGATGCTGTGTGCATATGCAATGCTAGCTTGATATCTCCA 745
Db 1 GAAGTGGCAAGCTTTTGTGCGATGCTGTGTATCATATGCAATGCTAGCTTGATATCTCCA 60

QY 746 TACAGGAGATCGTGATGCGGTCTCTACTTCCACATTTCACTTTATTTGAAGTA 805
Db 61 TACAGGAGATCGTGATGCGGTCTCTACTTCCAGATTTCTACTTTTGAAGTA 120

QY 806 TTTATTGATTTGCCCTTAAAGTTTGTGAAGCTCTGTATCTTAAGGCTTATACAGCTT 865
Db 121 TTTATTGATTTGCCCTTAAAGTTTGTGAAGCTCTGTATCTTAAGGCTTGTACAAGCTT 180

QY 866 GCACCTCAGCAAGAAAGATTAAAGTTTCACTGGAATTGATGATCCATACGAACCAATTT 925
Db 181 GCACCTCAGCAAGAAAGATTAAAGTTTCACTGGAATTGATGATCCATACGAACCAATTT 240

QY 926 AATGTTGAGATAGTAATTAAGATGAAGATGAGGAATGCCCTTCAACCAAGCAATGGCC 985
Db 241 AATGTTGAGATAGTAATTAAGATGAAGATGAGGAATGCCCTTCAACCAAGCAATGGCC 300

QY 986 AAGCAAGTTCTATGCTACCTTGAAGAAACCGATATTTGCAAGCTTAGTATATGATTTT 1045
Db 301 AAGCAAGTTCTATGCTACCTTGAAGAAACCGATATTTGCAAGCTTAGTATATGATTTT 357

QY 1046 GAGAAGATTGATCTGATTTCTTGTGTCATTACTTGTGGACACAATAAGATCTGTTGT 1105
Db 358 GAGAATATTGATCTGATTTCTTGTGTCATTACTTGTGGACACAATAAGATCTGTTGT-- 415

QY 1106 GGTCAATGAATAAAGGCATCAACATGTAGGAAGTAACAGAAAGTACGGTTTCATTCAGA 1165
Db 416 ---GCCATGAATAAAGGCATCAACATGTAGGAAGTAACAGAAAGTACGGTTTCATTCAGA 472

QY 1166 AACGATATGATTCATTCGTTTAA 1190
Db 473 ACGGAATACAGATTCATTCGTTTCA 497
```

Search completed: September 9, 2005, 12:40:22

Job time : 3862 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2005, 10:07:00 ; Search time 121 Seconds
(without alignments)
1096.354 Million cell updates/sec

Title: US-10-829-432-4
Perfect score: 1826
Sequence: 1 RPFHFINQTEPLVTHQPPP.....PKMAKQVLCYLENGYLQA 343

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1826	100.0	343	AAV44789	Corn Aden
2	887	48.6	224	AAV44788	Corn Aden
3	858	47.0	246	AAV44792	Wheat Ade
4	852	46.7	311	AAV44791	Soybean A
5	834	45.7	208	AAG35788	Arabidops
6	817	44.7	259	AAG14477	Arabidops
7	817	44.7	272	AAG14476	Arabidops
8	816	44.7	251	AAG38348	Arabidops
9	816	44.7	251	AAG11017	Arabidops
10	816	44.7	263	AAG38347	Arabidops
11	816	44.7	263	AAG11016	Arabidops
12	816	44.7	276	AAG38346	Arabidops
13	816	44.7	276	AAG11015	Arabidops
14	816	44.7	276	ADN74375	Thale cre
15	814	44.6	236	AAG14478	Arabidops
16	809.5	44.3	237	AAG53120	Arabidops
17	777.5	42.6	293	AAV77957	Arabidops
18	777.5	42.6	293	AAG29425	Arabidops
19	757.5	41.5	252	AAG47458	Arabidops
20	757.5	41.5	305	AAG47457	Arabidops
21	757.5	41.5	310	AAG47456	Arabidops
22	752.5	41.2	252	AAG14580	Arabidops
23	752.5	41.2	305	AAG14579	Arabidops
24	752.5	41.2	310	AAG14578	Arabidops
25	617.5	33.8	161	AAG53456	Arabidops

ALIGNMENTS

RESULT 1

AAV44789
ID AAV44789 standard; protein; 343 AA.

XX AC AAV44789;

DT 04-MAY-2000 (first entry)

DE Corn Adenylsulphate kinase-2.

XX Adenylsulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
KW 3'-Phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
KW corn; clone p0016.ctscj40rb; transgenic plant; screen; antibody.
XX Zea mays.

XX WO200004165-A1.

XX PD 27-JAN-2000.

XX PF 13-JUL-1999; 99WO-US015809.

XX PR 14-JUL-1998; 98US-0092833P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX PI Falco SC, Allen SM, Anderson SL;

XX DR WPI; 2000-182430/16.

XX N-PSDB; AAZ50160.

PT New nucleic acid molecule and chimeric gene encoding an adenosine-5'-
PT phosphosulphate kinase, useful for altering expression of sulfate
PT assimilation protein in plants.

XX Claim 6; Page 31-32; 42pp; English.

CC The present sequence is the corn adenylsulphate kinase (APS kinase),
CC also known as adenosine-5' phosphosulphate kinase. This is obtained from
CC clone p0016.ctscj40rb, derived from corn pooled tassels shoots, p0016 cDNA
CC library. This has 52% sequence identity to Catharanthus roseus APS
CC kinase. APS kinase is a sulphate assimilation protein, that catalyses the
CC conversion of adenosine-5' phosphosulphate (APS) to 3'-phospho- adenosine
CC -5' phosphosulphate (PAPS). The nucleotide sequence is used as probes and
CC primers to identify, obtain and synthesize sulphate assimilation proteins
CC from other plants. It is also used to produce transgenic plants, that are
CC useful for altering the expression levels of a sulphate assimilation
CC protein. The APS kinase peptides are useful for producing antibodies,

Abu23720 Protein e
Ada28651 Bacterial
Ada28185 Bacterial
Abu43064 Protein e
Abp39672 Staphyloc
Ada805772 Staphyloc
Ada27651 Bacterial
Adf06882 Bacterial
Abu18587 Protein e
Abu41025 Protein e
Abu49802 Protein e
Ada43993 Bacterial
Adn23471 Bacterial
Adc97735 Rhiobito
Abu31160 Protein e
Abu63457 Klebsiell
Abu67698 Phototrab
Ada28567 Bacterial
Ada21702 Bacterial
Abu48145 Protein e

26 571 31.3 200 6 ABU23720
27 555 30.4 199 8 ADS28651
28 534.5 29.3 202 8 ADS28185
29 521.5 28.6 199 6 ABU43064
30 521.5 28.6 204 5 ABP39672
31 521.5 28.6 204 8 ADS05772
32 519.5 28.5 196 8 ADS27651
33 517 28.3 226 7 ADF06882
34 514.5 28.2 197 6 ABU18587
35 514.5 28.2 204 6 ABU41025
36 510 27.9 213 6 ABU49802
37 505.5 27.7 199 6 ABU43993
38 505 27.7 652 8 ADN23471
39 504.5 27.6 627 7 ADC97735
40 503 27.5 201 6 ABU31160
41 503 27.5 216 7 ABU63457
42 501 27.4 211 6 ABU67698
43 498.5 27.3 660 8 ADS28567
44 496.5 27.2 621 8 ADS21702
45 495 27.1 201 6 ABU48145

```

CC that are used to screen and isolate cDNA clones
XX
SQ Sequence 343 AA;

Query Match      100.0%; Score 1826; DB 3; Length 343;
Best Local Similarity 100.0%; Pred. No. 2e-171;
Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 REPHEINQTEPLVTHTQPPSPAPGASQGGQGGNTLLSPTPTLAVILVNPORAPPVLP 60
DB 1 REPHEINQTEPLVTHTQPPSPAPGASQGGQGGNTLLSPTPTLAVILVNPORAPPVLP 60

QY 61 LTPSDAPLPALVIHGLTPRSSHSSAGLASDSGRREGEGRGARTHCHRGIGRWVRRRRNG 120
DB 61 LTPSDAPLPALVIHGLTPRSSHSSAGLASDSGRREGEGRGARTHCHRGIGRWVRRRRNG 120

QY 121 AAPGAPHSVPVKEPVMNSIKSTNHLWNCLIGQSDRQKLLGQKGCVVWITGLSGSGKS 180
DB 121 AAPGAPHSVPVKEPVMNSIKSTNHLWNCLIGQSDRQKLLGQKGCVVWITGLSGSGKS 180

QY 181 TLACALRELHCRGHLTYVLDGDLNRHGLNRDLSFKAEDRAENIRRVGEVAKLPADAGVI 240
DB 181 TLACALRELHCRGHLTYVLDGDLNRHGLNRDLSFKAEDRAENIRRVGEVAKLPADAGVI 240

QY 241 CIASLISPYRRDRDACCRRALLPHSNFIEVFIDLPLKICEARDPKGLYKLARTGKIKGFTGI 300
DB 241 CIASLISPYRRDRDACCRRALLPHSNFIEVFIDLPLKICEARDPKGLYKLARTGKIKGFTGI 300

QY 301 DDPEPPPINGEIVIMKDECPSPKAMAKQVLCYLENGYLOA 343
DB 301 DDPEPPPINGEIVIMKDECPSPKAMAKQVLCYLENGYLOA 343

RESULT 2
AAY44788
ID AAY44788 standard; protein; 224 AA.
XX
AC AAY44788;
XX
DT 04-MAY-2000 (first entry)
XX
DE Corn Adenylsulphate kinase-1.
XX
KW Adenylsulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
KW 3'-phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
KW corn; clone cen3n.pk0088.b10; transgenic plant; screen; antibody.
XX
OS Zea mays.
XX
PN WO200004165-A1.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015809.
XX
PR 14-JUL-1998; 98US-0092833P.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Falco SC, Allen SM, Anderson SL;
XX
DR WPI; 2000-182430/16.
DR N-PSDB; AA250159.
XX
PT New nucleic acid molecule and chimeric gene encoding an adenosine-5'
PT phosphosulfate kinase, useful for altering expression of sulfate
PT assimilation protein in plants.
XX
PS Claim 6; Page 29-30; 42pp; English.
XX
CC The present sequence is the corn adenylsulphate kinase (APS kinase),
CC also known as adenosine-5' phosphosulphate kinase. It is obtained from
CC clone cen3n.pk0088.b10, derived from corn endosperm, cen3n cDNA library.

```

```

CC This has 70% sequence identity to Catharanthus roseus APS kinase. APS
CC kinase is a sulphate assimilation protein, that catalyses the conversion
CC of adenosine-5' phosphosulphate (APS) to 3'-Phospho-adenosine-5'
CC phosphosulphate (PAPS). The nucleotide sequence is used as probes and
CC primers to identify, obtain and synthesise sulphate assimilation proteins
CC from other plants. It is also used to produce transgenic plants that are
CC useful for altering the expression levels of a sulphate assimilation
CC protein. The APS kinase peptides are useful for producing antibodies,
CC that are used to screen and isolate cDNA clones
XX
SQ Sequence 224 AA;

Query Match      48.6%; Score 887; DB 3; Length 224;
Best Local Similarity 77.0%; Pred. No. 7.4e-79;
Matches 171; Conservative 16; Mismatches 35; Indels 0; Gaps 0;

QY 121 AAPGEAPHSPVKEPVMNSIKSTNHLWNCLIGQSDRQKLLGQKGCVVWITGLSGSGKS 180
DB 2 AAAAVAGISSSSSALVTSTVGKSTNHLWHECAIQKERQGLLNQKGCVVWITGLSGSGKS 61

QY 181 TLACALRELHCRGHLTYVLDGDLNRHGLNRDLSFKAEDRAENIRRVGEVAKLPADAGVI 240
DB 62 TLACALRELHCRGHLTYVLDGDLNRHGLNRDLSFKAEDRAENIRRVGEVAKLPADAGLV 121

QY 241 CIASLISPYRRDRDACCRRALLPHSNFIEVFIDLPLKICEARDPKGLYKLARTGKIKGFTGI 300
DB 122 CIASLISPYRRDRDACCRRALLPHSNFIEVFIDLPLQVCEARDPKGLYKLARAGIKGFTGI 181

QY 301 DDPEPPPINGEIVIMKDECPSPKAMAKQVLCYLENGYLO 342
DB 182 DDPEPPSDCEIVIQCKVDCPSPESMAGHVVSYLETNGFLQ 223

RESULT 3
AAY44792
ID AAY44792 standard; protein; 246 AA.
XX
AC AAY44792;
XX
DT 04-MAY-2000 (first entry)
XX
DE Wheat Adenylsulphate kinase-1.
XX
KW Adenylsulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
KW 3'-phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
KW wheat; clone wrl.pk0101.e2; transgenic plant; screen; antibody.
XX
OS Triticum aestivum.
XX
PN WO200004165-A1.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US015809.
XX
PR 14-JUL-1998; 98US-0092833P.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Falco SC, Allen SM, Anderson SL;
XX
DR WPI; 2000-182430/16.
DR N-PSDB; AA250163.
XX
PT New nucleic acid molecule and chimeric gene encoding an adenosine-5'
PT phosphosulfate kinase, useful for altering expression of sulfate
PT assimilation protein in plants.
XX
PS Claim 6; Page 35; 42pp; English.
XX
CC The present sequence is the wheat adenylsulphate kinase (APS kinase),
CC also known as adenosine-5' phosphosulphate kinase. This is obtained from
CC clone wrl.pk0101.e2, derived from 7 day old light grown wheat root

```


PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-01324256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0137282P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 01-JUL-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.

PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159323P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.

```
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161404P.
PR 22-OCT-1999; 99US-0161405P.
PR 22-OCT-1999; 99US-0161406P.
PR 22-OCT-1999; 99US-0161359P.
PR 22-OCT-1999; 99US-0161360P.
PR 22-OCT-1999; 99US-0161361P.
PR 22-OCT-1999; 99US-0161920P.
PR 22-OCT-1999; 99US-0161922P.
PR 22-OCT-1999; 99US-0161932P.
PR 22-OCT-1999; 99US-0162142P.
PR 22-OCT-1999; 99US-0160768P.
PR 22-OCT-1999; 99US-0160770P.
PR 22-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161404P.
PR 22-OCT-1999; 99US-0161405P.
PR 22-OCT-1999; 99US-0161406P.
PR 22-OCT-1999; 99US-0161359P.
PR 22-OCT-1999; 99US-0161360P.
PR 22-OCT-1999; 99US-0161361P.
PR 22-OCT-1999; 99US-0161920P.
PR 22-OCT-1999; 99US-0161922P.
PR 22-OCT-1999; 99US-0161932P.
PR 22-OCT-1999; 99US-0162142P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 23-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134258P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 23-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134258P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
```

```
Query Match 45.7%; Score 834; DB 3; Length 208;
Best Local Similarity 74.8%; Pred. No. 1.2e-73;
Matches 154; Conservative 28; Mismatches 24; Indels 0; Gaps 0;
QY 137 MSNIGKSTNLWHNCLIGSDROKLLGOKGVVWITGLSGSGKSTLACLSRELHCRGHL 196
DB 1 MSTVGNSTNIFWQESPIGTERQKLLNQKGVVWITGLSGSGKSTLACLSRELNRGKL 60
QY 197 TYVLDGDNLRHGLNRLSFAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDAC 256
DB 61 SYLDGDNLRHGLNRLSFAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDAC 120
QY 257 RALLPHSNFIEVDLPLKICEARDPKGLYKLARTGKIKGFTGIDDPYPPEPPINGEIVIKM 316
DB 121 REMIONSSFIEVPMNLSQLCEARDPKGLYKLARAGKIKGFTGIDDPYSPINCEIETLKE 180
QY 317 KDECPSPKAMAKQVLCYLENGYIQ 342
DB 181 KECECPSPAMABEVSYLEDKGLQ 206
```

```
RESULT 6
AAG14477
ID AAG14477 standard; protein; 259 AA.
XX
AC AAG14477;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 14355.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 29-MAR-1999; 99US-0126264P.
PR 01-APR-1999; 99US-0126785P.
PR 06-APR-1999; 99US-0127462P.
PR 08-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 19-APR-1999; 99US-0129845P.
PR 21-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
```

PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145081P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.	Query Match 44.7%; Score 817; DB 3; Length 259;		
PR	05-AUG-1999;	99US-0147192P.	Best Local Similarity 60.6%; Pred. No. 7.6e-72;		
PR	05-AUG-1999;	99US-0147260P.	Matches 163; Conservative 32; Mismatches 52; Indels 22; Gaps 4;		
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	06-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.	QY	87 LASDSGRREGGR-----GARTHCHRGIGRWVRRRRRNGAAGPGEAPHSPVKEK	134
PR	10-AUG-1999;	99US-0148171P.	Db	1 MASPKGLSDNSRSVVVVRACVSMDSQTSLSHNKNGSIPELKSING-----HTGQKQG	53
PR	11-AUG-1999;	99US-0148319P.	QY	135 PVMSNIGKSTNILWHNCILIGOSDRKLLGQKGCYVWITGLSGSGKSTIACALSRELHCRG	194
PR	12-AUG-1999;	99US-0148341P.	Db	54 P-LSTVGNSTNIKWHECSVEKVDQRLLDQKGCYVWITGLSGSGKSTIACALNQLYQKG	112
PR	13-AUG-1999;	99US-0148565P.	QY	195 HLTYYLDGDNLRHGLNRDLSPKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRD	254
PR	13-AUG-1999;	99US-0149368P.	Db	113 KLCYILDGDNVVRHGLNRDLSPKAEDRAENIRRVGEVAKLFADAGIICIASLISPYRTDRD	172
PR	16-AUG-1999;	99US-0149175P.	QY	255 ACRALLPHSNFIEVFDLPKICEARDPKGLYKLARTGKIKGFTGIDDPYPPINGEIV	314
PR	18-AUG-1999;	99US-0149426P.	Db	173 ACRNLLPEGDFEVFMDVPLSVCEARDPKGLYKLARAGIKGFTGIDDPYPPLNCE--I	230
PR	20-AUG-1999;	99US-0149722P.	QY	315 KMDEECPSPKAMAKQVLCYLEENGYLQA	343
PR	20-AUG-1999;	99US-0149723P.	Db	231 SLGREGGTSPIEMAEKVVGYLNDKGYLOA	259
PR	20-AUG-1999;	99US-0149929P.	RESULT 7		
PR	23-AUG-1999;	99US-0149902P.	AAG14476		
PR	23-AUG-1999;	99US-0149930P.	ID	AAG14476 standard; protein; 272 AA.	
PR	25-AUG-1999;	99US-0150566P.	XX	AAG14476;	
PR	26-AUG-1999;	99US-0150884P.	XX	17-OCT-2000 (first entry)	
PR	27-AUG-1999;	99US-0151065P.	XX	Arabidopsis thaliana protein fragment SEQ ID NO: 14354.	
PR	27-AUG-1999;	99US-0151066P.	XX	Arabidopsis thaliana	
PR	27-AUG-1999;	99US-0151080P.	XX	Protein identification; signal transduction pathway; metabolic pathway;	
PR	30-AUG-1999;	99US-0151303P.	XX	hybridisation assay; genetic mapping; gene expression control; promoter;	
PR	31-AUG-1999;	99US-0151438P.	XX	termination sequence.	
PR	01-SEP-1999;	99US-0151930P.	XX	Arabidopsis thaliana.	
PR	07-SEP-1999;	99US-0152363P.	XX	EP1033405-A2.	
PR	10-SEP-1999;	99US-0153070P.	XX	06-SEP-2000.	
PR	13-SEP-1999;	99US-0153758P.	XX	25-FEB-2000; 2000EP-00301439.	
PR	16-SEP-1999;	99US-0154039P.	XX	25-FEB-1999; 99US-0121825P.	
PR	20-SEP-1999;	99US-0154779P.	XX	05-MAR-1999; 99US-012180P.	
PR	22-SEP-1999;	99US-0155139P.	XX	09-MAR-1999; 99US-0123548P.	
PR	23-SEP-1999;	99US-0155486P.	XX		
PR	24-SEP-1999;	99US-0155659P.	XX		
PR	28-SEP-1999;	99US-0156458P.	XX		
PR	29-SEP-1999;	99US-0156596P.	XX		
PR	04-OCT-1999;	99US-0157117P.	XX		
PR	05-OCT-1999;	99US-0157753P.	XX		
PR	06-OCT-1999;	99US-0157865P.	XX		
PR	07-OCT-1999;	99US-0158029P.	XX		
PR	08-OCT-1999;	99US-0158232P.	XX		
PR	12-OCT-1999;	99US-0158369P.	XX		
PR	13-OCT-1999;	99US-0159293P.	XX		
PR	13-OCT-1999;	99US-0159294P.	XX		
PR	13-OCT-1999;	99US-0159295P.	XX		
PR	14-OCT-1999;	99US-0159329P.	XX		
PR	14-OCT-1999;	99US-0159330P.	XX		
PR	14-OCT-1999;	99US-0159331P.	XX		

PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139494P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144633P.
PR 21-JUL-1999; 99US-0144844P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.

PR	07-OCT-1999;	99US-0158029P.	PN	EP1033405-A2.
PR	08-OCT-1999;	99US-0158232P.	XX	
PR	12-OCT-1999;	99US-0158369P.	PD	06-SEP-2000.
PR	13-OCT-1999;	99US-0159293P.	XX	
PR	13-OCT-1999;	99US-0159294P.	PF	25-FEB-2000; 2000EP-00301439.
PR	13-OCT-1999;	99US-0159295P.	XX	
PR	14-OCT-1999;	99US-0159329P.	PR	25-FEB-1999;
PR	14-OCT-1999;	99US-0159330P.	PR	99US-0121825P.
PR	14-OCT-1999;	99US-0159331P.	PR	99US-0123180P.
PR	14-OCT-1999;	99US-0159637P.	PR	05-MAR-1999;
PR	14-OCT-1999;	99US-0159638P.	PR	99US-0123548P.
PR	18-OCT-1999;	99US-0159584P.	PR	23-MAR-1999;
PR	18-OCT-1999;	99US-0159585P.	PR	99US-0125788P.
PR	21-OCT-1999;	99US-0160741P.	PR	25-MAR-1999;
PR	21-OCT-1999;	99US-0160767P.	PR	99US-0126264P.
PR	21-OCT-1999;	99US-0160768P.	PR	29-MAR-1999;
PR	21-OCT-1999;	99US-0160770P.	PR	99US-0126785P.
PR	21-OCT-1999;	99US-0160814P.	PR	01-APR-1999;
PR	21-OCT-1999;	99US-0160815P.	PR	99US-0128234P.
PR	22-OCT-1999;	99US-0160980P.	PR	99US-0127462P.
PR	22-OCT-1999;	99US-0160981P.	PR	99US-0126785P.
PR	22-OCT-1999;	99US-0160989P.	PR	06-APR-1999;
PR	25-OCT-1999;	99US-0161404P.	PR	99US-0128714P.
PR	25-OCT-1999;	99US-0161405P.	PR	99US-0129845P.
PR	25-OCT-1999;	99US-0161406P.	PR	99US-0130077P.
PR	26-OCT-1999;	99US-0161359P.	PR	99US-0130449P.
PR	26-OCT-1999;	99US-0161360P.	PR	99US-0130510P.
PR	26-OCT-1999;	99US-0161361P.	PR	99US-0130891P.
PR	28-OCT-1999;	99US-0161920P.	PR	99US-0131449P.
PR	28-OCT-1999;	99US-0161992P.	PR	99US-0132048P.
PR	28-OCT-1999;	99US-0161993P.	PR	99US-0132407P.
PR	29-OCT-1999;	99US-0162142P.	PR	99US-0132484P.
Query Match 44.7%; Score 817; DB 3; Length 272;				
Best Local Similarity 60.6%; Pred. No. 8.1e-72;				
Matches 163; Conservative 32; Mismatches 52; Indels 22; Gaps 4;				
QY	87	LASDGRREGGR-----GARTHCHRGIGRWVRRRRNGAAGPEAPHSFVKEK 134		
Db	14	MASPKGLSDNSRSVVVACVSMDSQTLSHNKGSIPELKSING-----HTGQKG 66		
QY	135	PWNSNGKSTNLIWNCLIGQSDRQKLQKQGVVWTLGSGSKSTLACALSRELHCRG 194		
Db	67	P-LSTVGNSTNFKWHECSVEKVDQRLLDQKGVVWTLGSGSKSTLACALNQMLYQK 125		
QY	195	HLTYVLGDNLRHGLNRDLSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRD 254		
Db	126	KLCYILDGDNVRHGLNRDLSFKAEDRAENIRRVGEVAKLFADAGIICIASLISPYRTDRD 185		
QY	255	ACRALPHSNFIEVFIDLPLKICEARDPKGLYKLARTGKIKGFTGIDDPYEPPIGEIVI 314		
Db	186	ACRNLLPEGDFVEVFMDFVLSVCEARDPKGLYKLARAGIKGFTGIDDPYEPPLNCE--I 243		
QY	315	KMKDECPSPKMAKQVLCYLSENGYLQA 343		
Db	244	SLGREGGTSPIEMAEKVGVLDNKGYLQA 272		
RESULT 8				
AAG38348				
ID	AAG38348 standard; protein; 251 AA.			
XX				
AC	AAG38348;			
XX				
DT	18-OCT-2000 (first entry)			
XX				
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 47297.			
XX				
KW	Protein identification; signal transduction pathway; metabolic pathway;			
KW	hybridisation assay; genetic mapping; gene expression control; promoter;			
KW	termination sequence.			
XX				
OS	Arabidopsis thaliana.			
XX				

XX Arabidopsis thaliana protein fragment SEQ ID NO: 9573.
DE Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS EP1033405-A2.
PN 06-SEP-2000.
PD 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 21-MAY-1999; 99US-0135629P.
PR 24-MAY-1999; 99US-0136021P.
PR 25-MAY-1999; 99US-0136392P.
PR 27-MAY-1999; 99US-0136782P.
PR 28-MAY-1999; 99US-0137222P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 20-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149923P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.

PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-01511303P.
PR	31-AUG-1999;	99US-01511438P.
PR	01-SEP-1999;	99US-01511930P.
PR	07-SEP-1999;	99US-0152163P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155113P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 44.7%; Score 816; DB 3; Length 251;		
Best Local Similarity 64.4%; Pred. No. 9.1e-72;		
Matches 159; Conservative 32; Mismatches 40; Indels 16; Gaps 4;		
QY	100	GARTCHRGIGRWRRRRNGAAGEAP---HSPVKEKPVMSNICKSTNHLWNCLIGOS 156
Db	18	GSQTLSH-----NKNGSIPEVKSINGHTGQKQP-ISTGVNSTNKKHCEKSVKV 66
QY	157	DRQKLLGQKGVWITGLSGSKSTLACALSRELHCRGHLTYVLDGDNLRHGLNRDLSFK 216
Db	67	DRQRLDQKGVWITGLSGSKSTLACALNQMLYQKGLCYILDDGNVRHGLNRDLSFK 126
QY	217	AEDRAENTRRVEGAKLFADAGVICTASLISPYRRDRACRALLPHSNFIEVFDLPKI 276
Db	127	AEDRAENTRRVEGAKLFADAGIICIASLISPYRTDRACRSLLEGGDFVEYFMDVPLSV 186
QY	277	CEARDPKGLYKLARTKGFTGIDDPVEPPINGEIVIMKMDKECPSPKAMAKOVLVLYE 336
Db	187	CEARDPKGLYKLARAGKINGFTGIDDPVEPLNCE--ISLREGGTSPTEMAEKVVGYLD 244
QY	337	ENGYLQA 343
Db	245	NGGYLQA 251

RESULT 10
AAG38347
ID AAG38347 standard; protein; 263 AA.
XX
AC AAG38347;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47296.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 200EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.

PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 44.7%; Score 816; DB 3; Length 263;		
Best Local Similarity 64.4%; Pred. No. 9.7e-72;		
Matches 159; Conservative 32; Mismatches 40; Indels 16; Gaps 4;		
QY	100	GARTHCHRGIGRWRRRRRNGAAGEAP---HSPVKEKPMNSGNKSTWILWHNCLIGOS 156

Db	30	GSQTLSH-----NKGSIPEVKSINGHTCQKQCP-LSTVGNSTNIKWHCESVKV 78
QY	157	DRKLLGQKGVVWITGLSGSGKSTLACALSREILHCRGHLTYVLDGDNLRHGLNRLDSFK 216
Db	79	DRQRLDQKGVVWITGLSGSGKSTLACALNQMLYQKGKLCYIIDDGNVRHGLNRLDSFK 138
QY	217	AEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDACRALLPHSNFIEVFIDLPLKI 276
Db	139	AEDRAENIRRVGEVAKLFADAGIICIASLISPYRTDRDACRSLLPEDGFEVFDVPLSV 198
QY	277	CEARDPKGLYKLGATGKIKGFTGIDDPYEPPIGEIIVKMDRECPSPKAKQVLCYLE 336
Db	199	CEARDPKGLYKLGARAGIKGFTGIDDPYEPPLNCE--ISLREGGTSPIENAEKVVGILD 256
QY	337	ENGYLOA 343
Db	257	NKGYLOA 263
RESULT 12		
AAG38346		
ID	AAG38346	standard; protein; 276 AA.
AC	AAG38346;	
XX	18-OCT-2000	(first entry)
DE	Arabidopsis thaliana	protein fragment SEQ ID NO: 47295.
KW	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
XX	termination sequence.	
OS	Arabidopsis thaliana.	
PN	EP1033405-A2.	
PD	06-SEP-2000.	
PF	25-FEB-2000; 2000EP-00301439.	
PR	25-FEB-1999;	99US-0121825P.
PR	05-MAR-1999;	99US-0123180P.
PR	09-MAR-1999;	99US-0123548P.
PR	23-MAR-1999;	99US-0125788P.
PR	25-MAR-1999;	99US-0126264P.
PR	29-MAR-1999;	99US-0126785P.
PR	01-APR-1999;	99US-0127462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	23-APR-1999;	99US-0130891P.
PR	28-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	04-MAY-1999;	99US-0132407P.
PR	05-MAY-1999;	99US-0132484P.
PR	06-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	07-MAY-1999;	99US-0132487P.
PR	11-MAY-1999;	99US-0132863P.
PR	14-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	18-MAY-1999;	99US-0134370P.
PR	19-MAY-1999;	99US-0134768P.
PR	20-MAY-1999;	99US-0134941P.
PR	21-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151338P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.

```
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 44.7%; Score 816; DB 3; Length 276;
Best Local Similarity 64.4%; Pred. No. 1e-71;
Matches 159; Conservative 32; Mismatches 40; Indels 16; Gaps 4;

QY 100 GARTCHRGIGRVRRRRNGAAGPGEAP---HSPVKEKPEVMSNICKSTNHLWHNCLIGOS 156
Db 43 GSQTLSH-----NKGSIPEVKSINGHTGOKGP-LSTGVNSTNKHWCSEVKV 91

QY 157 DRQKLLGQKGVWITGLSGSGKSTLACALSRELHCRGHLTYVLQDNLRLHGLNRDLSPK 216
Db 92 DRQRLDDQKGVWITGLSGSGKSTLACALNQLKQKGLCYLLDGNVRLHGLNRDLSPK 151

QY 217 AEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDACRALLPHSNFIEVFDLPK 276
Db 152 AEDRAENIRRVGEVAKLFADAGIICIASLISPYRTDRDACRSLLPEDGFVEFMDVPLSV 211

QY 277 CEAROPKGLYKLARTGKIKGFTGIDDPYEPPIGEIVIKMKDEECPSKAMAKOVLVLYLE 336
Db 212 CEAROPKGLYKLARAKIKGFTGIDDPYEPPLNCE--ISLREGGTSPTIEMAERKVVGYLD 269

QY 337 ENGYLQA 343
Db 270 NKGYLQA 276

RESULT 13
AAG11015
ID AAG11015 standard; protein; 276 AA.
XX AC AAG11015;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 9571.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 08-APR-1999; 99US-0128714P.
XX PR 16-APR-1999; 99US-0129845P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 21-APR-1999; 99US-0130449P.
XX PR 23-APR-1999; 99US-0130510P.
XX PR 23-APR-1999; 99US-0130891P.
XX PR 28-APR-1999; 99US-0131449P.
XX PR 30-APR-1999; 99US-0132048P.
XX PR 04-MAY-1999; 99US-0132407P.
XX PR 05-MAY-1999; 99US-0132484P.
XX PR 06-MAY-1999; 99US-0132485P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 07-MAY-1999; 99US-0132487P.
XX PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137232P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
```



```
XX PS Claim 1; SEQ ID NO 2270; 134pp; English.
XX CC This invention relates to a novel method for altering one or more plant
XX CC characteristics. Specifically, it refers to identifying genes that are up
XX CC - or down-regulated in transgenic plants overexpressing the heterodimeric
XX CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
XX CC alter plant characteristics accordingly. The present invention describes
XX CC generating transgenic plants for the production of growth regulators,
XX CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
XX CC the altered plant characteristics are selected from increased yield or
XX CC biomass, enhanced survival capacity, stress tolerance, plant architecture
XX CC or physiology, altered endoreduplication, biochemistry, signal
XX CC transduction, storage lipid mobilization and/or altered photosynthesis,
XX CC each relative to the corresponding wild type plants. Accordingly, these
XX CC sequences can also be useful as positive or negative selectable markers
XX CC during transformation of cells or tissues. The identified genes play a
XX CC role in a variety of biological processes such as DNA replication, cell
XX CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
XX CC transcription factors. This polypeptide sequence is thale cress protein
XX CC expressed by a gene repressed 1.3 fold or more in plants overexpressing
XX CC the E2Fa/Dpa transcription factor, given in an exemplification of the
XX CC invention.
XX SQ Sequence 276 AA;
Query Match 44.7%; Score 816; DB 8; Length 276;
Best Local Similarity 64.4%; Pred. No. 1e-71;
Matches 159; Conservative 32; Mismatches 40; Indels 16; Gaps 4;
QY 100 GARTCHRCIGRWRRRRNGAAGEAP---HSPVKEKPVMSNICKSTNIMHNLIGQS 156
Db 43 GSQTLSH-----NKGSIPEVKSI NGHTGQKQGP-LSTGVGNTNIKWHCEKVEKV 91
QY 157 DRQKLLGQKGVVWVTGLSGSGKSTLACALSELHCRGHLYTVLDGDLNRLHGLNRLDSFK 216
Db 92 DRQRLDQKGVVWVTGLSGSGKSTLACALNQMLYQKGLCYLVDGDNVNRHGLNRLDSFK 151
QY 217 AEDRAENIRRVGEVAKLFADAGVICIASLISPYRDRDRCALLPHSNFIEVFDLPKI 276
Db 152 AEDRAENIRRVGEVAKLFADAGIICIASLISPYRDRDRCALLPEGDFVEFMDVPLSV 211
QY 277 CEARDPKGLYKLARTGKIKGFTGIDDPYEPPIINGEIVIMKDECECPSPKAMAKQVLCYLE 336
Db 212 CEARDPKGLYKLARAGKIKGFTGIDDPYEPPLNCE--ISLREGCGTSPTEMAEKVVGYLD 269
QY 337 ENGYLQA 343
Db 270 NRGYLQA 276
RESULT 15
AAG14478
ID AAG14478 standard; protein; 236 AA.
XX AC AAG14478;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 14356.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX
```


PR 13-JUL-1999;	99US-0143542P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 21-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160982P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0160989P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161404P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145919P.	PR 25-OCT-1999;	99US-0161406P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161359P.
PR 02-AUG-1999;	99US-0145386P.	PR 26-OCT-1999;	99US-0161360P.
PR 02-AUG-1999;	99US-0146388P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161920P.
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161992P.
PR 04-AUG-1999;	99US-0147204P.	PR 28-OCT-1999;	99US-0161993P.
PR 04-AUG-1999;	99US-0147302P.	PR 29-OCT-1999;	99US-0161993P.
PR 05-AUG-1999;	99US-0147192P.	PR 29-OCT-1999;	99US-0162142P.
PR 05-AUG-1999;	99US-0147260P.		
PR 06-AUG-1999;	99US-0147303P.		
PR 06-AUG-1999;	99US-0147416P.		
PR 09-AUG-1999;	99US-0147493P.		
PR 09-AUG-1999;	99US-0147935P.		
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.		
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148565P.		
PR 13-AUG-1999;	99US-0148684P.		
PR 16-AUG-1999;	99US-0149368P.		
PR 17-AUG-1999;	99US-0149175P.		
PR 18-AUG-1999;	99US-0149426P.		
PR 20-AUG-1999;	99US-0149722P.		
PR 20-AUG-1999;	99US-0149723P.		
PR 20-AUG-1999;	99US-0149929P.		
PR 23-AUG-1999;	99US-0149902P.		
PR 23-AUG-1999;	99US-0149930P.		
PR 25-AUG-1999;	99US-0150566P.		
PR 26-AUG-1999;	99US-0150884P.		
PR 27-AUG-1999;	99US-0151065P.		
PR 27-AUG-1999;	99US-0151066P.		
PR 27-AUG-1999;	99US-0151080P.		
PR 30-AUG-1999;	99US-0151303P.		
PR 31-AUG-1999;	99US-0151438P.		
PR 01-SEP-1999;	99US-0151930P.		
PR 07-SEP-1999;	99US-0152363P.		
PR 10-SEP-1999;	99US-0153070P.		
PR 13-SEP-1999;	99US-0153758P.		
PR 15-SEP-1999;	99US-0154018P.		
PR 16-SEP-1999;	99US-0154039P.		
PR 20-SEP-1999;	99US-0154779P.		
PR 22-SEP-1999;	99US-0155139P.		
PR 23-SEP-1999;	99US-0155486P.		
PR 24-SEP-1999;	99US-0155659P.		
PR 28-SEP-1999;	99US-0156458P.		
PR 29-SEP-1999;	99US-0156596P.		

Query Match 44.6%; Score 814; DB 3; Length 236;
Beat Local Similarity 65.2%; Pred. No. 1.3e-71;
Matches 159; Conservative 30; Mismatches 45; Indels 10; Gaps 3;

Qy	100	GARTHCHRGIGRWVRRRRRNGAAPCEAPHSPVKEKPVMSNIGKSTNIIWHNCLIGQSDRQ	159
Db	3	GSQTLSHNKNGSIPELKSING-----HTQKQGP-LSTVGNSTNIKWHECSVEKVDKQ	54
Qy	160	KLLGQKGVWVITGLSGSGKSTLACALSRELHCRGHLTYVLDGDNLRHGLNRDLSFKAED	219
Db	55	RLLDQKGVIVWVITGLSGSGKSTLACALNQMLYQKGLCYILDGDNVRHGLNRDLSFKAED	114
Qy	220	RAENIRRVGEVAKLFADAGVTCIASLISPYRRDRDACLPHSNFIEVFDLPKICEA	279
Db	115	RAENIRRVGEVAKLFADAGIICIASLISPYRTDRDACLNRLLPEGDFVEFMDVPLSVCEA	174
Qy	280	RDPKGLYKARTGKIKGFTGIDDPYEPPIGEIVIKMKDECPSPKAMAKQVLCYLEENG	339
Db	175	RDPKGLYKARAGIKGFTGIDDPYEPPLNCE--ISLREGGTSPIEMAERKVVGYLDNKG	232
Qy	340	YLQA	343
Db	233	YLQA	236

Search completed: September 9, 2005, 13:00:51
Job time : 123 secs

THIS PAGE BLANK (USPTO)


```

; TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE:  2709.1002-001
; CURRENT APPLICATION NUMBER:  US/09/543.681A
; CURRENT FILING DATE:  2000-04-05
; PRIOR APPLICATION NUMBER:  US 60/128,706
; PRIOR FILING DATE:  1999-04-09
; NUMBER OF SEQ ID NOS:  8344
; SEQ ID NO 7167
; LENGTH:  226
; TYPE:  PRT
; ORGANISM:  Proteus mirabilis
US-09-543-681A-7167

Query Match      28.3%; Score 517; DB 4; Length 226;
Best Local Similarity  47.1%; Pred. No. 1.6e-44;
Matches 107; Conservative 31; Mismatches 57; Indels 32; Gaps 4;

QY      133  EKPWMSNIGKS-----TWLWHNCILIGSDRQKLLGQKGVVWITGLSGSGKST 181
DB      5  DPPFSAIGRKRLTGKRVITHQIDVHWHQIIGLKEREAQQVHGKCVLWFTGLSGSGKST 64

QY      182  LACALSRELHCRGHL-----TYVLDDGNLRHGLNRDLSFKAEDRAENIRRVGEVAKLPAD 236
DB      65  LADALEQLYQYSTLHAPIRTYLLDGNLRHGLCHDLGFSQDRHENIRRVGEVAKLMVD 124

QY      237  AGVICIASLISPYRRDRDACLALLPHSNFIEVFDLPLKICEARDPKGLYKLARTGKIG 296
DB      125  AGLVLVTAFLISPYQODRQVRERFAQGRFIEFVDTPLALCEARDPKGLYOKARRGEIKQ 184

QY      297  FTGIDDPYEPP-----INGEIVIKMKDECPSPKAMAKQVLCYLEE 337
DB      185  FSGIDSPYEPTPEIHLDSLAI-----NELTQQILAVLQQ 221

RESULT 3
US-09-489-039A-9974
; Sequence 9974, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT:  Gary Breton et. al
; TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE:  2709.2004001
; CURRENT APPLICATION NUMBER:  US/09/489,039A
; CURRENT FILING DATE:  2000-01-27
; PRIOR APPLICATION NUMBER:  US 60/117,747
; PRIOR FILING DATE:  1999-01-29
; NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 9974
; LENGTH:  216
; TYPE:  PRT
; ORGANISM:  Klebsiella pneumoniae
US-09-489-039A-9974

Query Match      27.5%; Score 503; DB 4; Length 216;
Best Local Similarity  57.7%; Pred. No. 4e-43;
Matches 97; Conservative 21; Mismatches 50; Indels 0; Gaps 0;

QY      145  NILWHNCILIGSDRQKLLGQKGVVWITGLSGSGKSTLACALSRELHCRGHLTYVLDDGN 204
DB      22  NVWHAHPVTQQREQHHGRGVVWITGLSGSGKSVAGALEEALHERGVSTYLLDGN 81

QY      205  LRHGLNRDLSFKAEDRAENIRRVGEVAKLPADAGVICIASLISPYRRDRDACLALLPHSN 264
DB      82  VRHGLCSDLGFSDEDRKENIRRVGEVARLWMDAGLVLTAFISPHRAERQMVRRLEGGR 141

QY      265  FIEVFIDLPLKICARDPKGLYKLARTCKIKGFTGIDDPYEPPIGEI 312
DB      142  FIEVFDTPLAICARDPKGLYKKARAGELRNFTGIDSVYEAPEKAEI 189

RESULT 4
US-08-879-561-10

```

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Hawkins, Philip R.
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/879,561
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0325 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 624 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LUNGN0T02
CLONE: 373887
US-08-879-561-3

Query Match	27.1%	Score 495;	DB 2;	Length 624;
Best Local Similarity	46.1%;	Pred. No. 1.2e-41;		
Matches 106;	Conservative 38;	Mismatches 60;	Indels 26;	Gaps 6;

QY	125	EAPHSVPKPKPVMSN-----IGKSTNTLHNCNLIGQSDROKLLGQ-----KGCVVWITGLS	175
DB	2	EIGSLCKYKVLNNAQWGMQRATNTYYQAHVSRNKRQGVQVTRGPGFGCTVWLTGLS	61
QY	176	GSQKSTLACALSBELHCRGHLTVYVLDGDNLRHGLNRDLSPKAEDRAENIRRVGEVAKLFA	235
DB	62	GAGKTTVSMALBYLVCHGIPCYVTLTDGDNIRQGLKNLGFSPEDREENVRARIAEVAKLFA	121
QY	236	DAGVICIASLISPVRRDRDAGR-----ALLPHSNFIEVFIDLPLKICEARDPKGLYKLLAR	290
DB	122	DAGLVCTISFISPYQDRNNARQHEGASLP-----FFEVFVDAPLHVCEQRDVKGLYKCAR	178
QY	291	TGKIKGFTGIDDPYEPPIINGEIVIMKMD-----EBCSPSKAMAKQVLCYLEE	337
DB	179	AGBIKGTGIDSYEKPEAPELVLTDCSDVND-----VQOVVELLOE	222

RESULT 6
US-08-879-561-9
; Sequence 9, Application US/08879561
; Patent No. 5817482
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga

```

1 APPLICANT: Hillman, Jennifer L.
2 APPLICANT: Hawkins, Phillip R.
3 APPLICANT: Guegler, Karl J.
4 APPLICANT: Corley, Neil C.
5 TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES
6 NUMBER OF SEQUENCES: 12
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: Incyte Pharmaceuticals, Inc.
9 STREET: 3174 Porter Drive
10 CITY: Palo Alto
11 STATE: CA
12 COUNTRY: USA
13 ZIP: 94304
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Diskette
16 COMPUTER: IBM Compatible
17 OPERATING SYSTEM: DOS
18 SOFTWARE: FastSeq for Windows Version 2.0
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/879,561
21 FILING DATE: Herewith
22 CLASSIFICATION: 424
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER:
25 FILING DATE:
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Billings, Lucy J.
28 REGISTRATION NUMBER: 36,749
29 REFERENCE/DOCKET NUMBER: PF-0325 US
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 415-855-0555
32 TELEFAX: 415-845-4166
33 TELEX:
34 INFORMATION FOR SEQ ID NO: 9:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 624 amino acids
37 TYPE: amino acid
38 STRANDEDNESS: single
39 TOPOLOGY: linear
40 IMMEDIATE SOURCE:
41 LIBRARY: GenBank
42 CLONE: 1109676
43 US-08-879-561-9
44
45 Query Match 27.1%; Score 495; DB 2; Length 6
46 Best Local Similarity 46.1%; Pred. No. 1.2e-41;
47 Matches 106; Conservative 38; Mismatches 60; Indels
48
49 Qy 125 EAPHSPVKEKPVMSN-----TGSNTILMHNCLIGQSDRQKLJGQ-
50 Db 2 ETPGSLCKKVLSSNNAQNMGQMQRATNTVYQAHVSRNKRKGQVVGTR
51
52 Qy 176 GSGKSTLACASRELHCRGHLYTVLDGDNLRHGLNRDLSPKAEDRA
53 Db 62 GAGKTTVSMALSEYLVCHGIPCYTLDDGNIIRQGNLKNLGFSPEDRE
54
55 Qy 236 DAGVICIASLIPYRRDRDAC-----ALLPHSNFIEVFDLPKIKI
56 Db 122 DAGLVCIITSFIPYQDRNNAQHEGASLP---FFEVFDVADAPLHV
57
58 Qy 291 TGKIKGFTGIDDPYEPPIINGEIVTKMKD-----EECPSPKAMAKQVLCV
59 Db 179 AGEIKGFTGIDSEYEKPAPELVTKTDSQVNDK-----VOOVVE

```

RESULT 7
US-09-949-016-7339
; Sequence 7339, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION
; AND TREATMENT OF SUCH DISEASES

[illegible]

Search completed: September 9, 2005, 13:05:05
Job time : 53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2005, 12:58:46 ; Search time 112 Seconds
(without alignments)
1207.946 Million cell updates/sec

Title: US-10-829-432-4
Perfect score: 1826
Sequence: 1 RPFHINTEPLVTHTQPP.....PKMAKQVLCYLENGYLQA 343

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1777461 seqs, 394431504 residues

Total number of hits satisfying chosen parameters: 1777461

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
 - 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1826	100.0	343	16	US-10-829-432-4
2	1168.5	64.0	270	16	US-10-425-115-301616
3	1013.5	55.5	253	15	US-10-425-114-47513
4	972	53.2	207	16	US-10-739-930-10868
5	936	51.3	188	16	US-10-767-701-40866
6	903	49.5	129	16	US-10-425-115-247552
7	900	49.3	204	16	US-10-437-963-154262
8	891.5	48.8	274	15	US-10-425-114-41817
9	887	48.6	224	16	US-10-829-432-2
10	867	47.5	349	16	US-10-425-115-319790
11	858	47.0	246	16	US-10-829-432-10
					Sequence 4, Appli
					Sequence 301616, A
					Sequence 47513, A
					Sequence 10868, A
					Sequence 40866, A
					Sequence 247552,
					Sequence 154262,
					Sequence 41817, A
					Sequence 2, Appli
					Sequence 319790,
					Sequence 10, Appl

12	854	46.8	215	15	US-10-425-114-55578	Sequence 55578, A
13	854	46.8	265	15	US-10-424-599-205747	Sequence 205747,
14	853	46.7	207	15	US-10-424-599-262462	Sequence 262462,
15	852	46.7	311	16	US-10-829-432-8	Sequence 8, Appli
16	845	46.3	234	15	US-10-425-114-61644	Sequence 61644, A
17	839	45.9	246	15	US-10-425-114-47070	Sequence 47070, A
18	833	45.6	312	16	US-10-829-432-13	Sequence 13, Appl
19	830	45.5	246	15	US-10-425-114-66421	Sequence 66421, A
20	816	44.7	276	16	US-10-829-432-14	Sequence 14, Appl
21	777.5	42.6	293	14	US-10-342-224-66	Sequence 66, Appl
22	770	42.2	173	16	US-10-437-963-175387	Sequence 175387,
23	723.5	39.6	301	16	US-10-425-115-219699	Sequence 219699,
24	699	38.3	207	15	US-10-424-599-262463	Sequence 262463,
25	674.5	36.9	185	16	US-10-425-115-301613	Sequence 301613,
26	611	33.5	158	16	US-10-437-963-148349	Sequence 148349,
27	601.5	32.9	203	16	US-10-425-115-219698	Sequence 219698,
28	571	31.3	200	15	US-10-282-122A-51644	Sequence 51644, A
29	555	30.4	199	15	US-10-369-493-17684	Sequence 17684, A
30	534.5	29.3	202	15	US-10-369-493-17218	Sequence 17218, A
31	529.5	29.0	162	9	US-09-795-693-40	Sequence 40, Appl
32	529.5	29.0	162	14	US-10-156-239-40	Sequence 40, Appl
33	529.5	29.0	162	14	US-10-199-485-40	Sequence 40, Appl
34	521.5	28.6	199	15	US-10-282-122A-70988	Sequence 70988, A
35	521.5	28.6	204	18	US-10-724-972A-5057	Sequence 5067, Ap
36	519.5	28.5	196	15	US-10-369-493-16684	Sequence 16684, A
37	514.5	28.2	197	15	US-10-282-122A-46511	Sequence 46511, A
38	514.5	28.2	204	15	US-10-282-122A-68949	Sequence 68949, A
39	510	27.9	213	15	US-10-282-122A-77726	Sequence 77726, A
40	505.5	27.7	199	15	US-10-282-122A-71917	Sequence 71917, A
41	505	27.7	652	15	US-10-369-493-6124	Sequence 6124, Ap
42	503	27.5	201	15	US-10-282-122A-59084	Sequence 59084, A
43	502	27.5	127	16	US-10-767-701-61157	Sequence 61157, A
44	498.5	27.3	660	15	US-10-369-493-17600	Sequence 17600, A
45	496.5	27.2	621	15	US-10-369-493-10735	Sequence 10735, A

ALIGNMENTS

RESULT 1
US-10-829-432-4
; Sequence 4, Application US/10829432
; Publication No. US20040177401A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio
; APPLICANT: Allen, Stephen
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-B
; CURRENT APPLICATION NUMBER: US/10/829,432
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/720,384A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Zea mays
US-10-829-432-4

Query Match	100.0%	Score	1826;	DB	16;	Length	343;
Best Local Similarity	100.0%	Pred. No.	9.5e-154;				
Matches	343;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Oy	1	RPFHINTEPLVTHTQPPSPAPGASQGGQGNLTLLSPPTTLAVILVNFQAPPVLP	60				
Db	1	RPFHINTEPLVTHTQPPSPAPGASQGGQGNLTLLSPPTTLAVILVNFQAPPVLP	60				
Oy	61	LTPSDAPLPAIVHGLTPRSSHSSAGLASDSRGREGGARTHCHRGIGRWVRRRRNG	120				

```
Db 61 LTPSDAPLALVIHGLTPRSSHSSAGLASDSGRREGGARGARTHCHRGIGRWVRRRRNG 120
Qy 121 AAPGEAPHSPVKEKPVMSNIGKSTNHLWNCLIGOSDRQKLLGOKGCVVWITGLSGSGKS 180
Db 121 AAPGEAPHSPVKEKPVMSNIGKSTNHLWNCLIGOSDRQKLLGOKGCVVWITGLSGSGKS 180
Qy 181 TLACALSRELHCRGHLTYVLDGDNLRHGLNRLDLSFKAEDRAENIRRVGEVAKLFADAGVI 240
Db 181 TLACALSRELHCRGHLTYVLDGDNLRHGLNRLDLSFKAEDRAENIRRVGEVAKLFADAGVI 240
Qy 241 CIASLISPYRRDRDACRALLPHSNFIEVDLPLKICEARDPKGLYKLARTGKIKGFTGI 300
Db 241 CIASLISPYRRDRDACRALLPHSNFIEVDLPLKICEARDPKGLYKLARTGKIKGFTGI 300
Qy 301 DDPEYPPINGEIVIKMKDEECPSPKAMAKQVLCYLEENGYLQA 343
Db 301 DDPEYPPINGEIVIKMKDEECPSPKAMAKQVLCYLEENGYLQA 343

RESULT 2
US-10-425-115-301616
; Sequence 301616, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 301616
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MR74577_38151C.1.ppep
US-10-425-115-301616

Query Match 64.0%; Score 1168.5; DB 16; Length 270;
Best Local Similarity 84.5%; Pred. No. 2.1e-95;
Matches 235; Conservative 6; Mismatches 28; Indels 9; Gaps 4;

Qy 66 APLPALVIHGLTPRSSHSSAGLASDSGRREGGARGARTHCHRGIGRWVRRRRNGAAPGE 125
Db 2 ASLP--VPHTL--PRASPAIVGAA--RGRAAVRVRTATATGALGGCGGGGMEQRH---GE 52

Qy 126 APSPVKEKPVMSNIGKSTNHLWNCLIGOSDRQKLLGOKGCVVWITGLSGSGKSTLACA 185
Db 53 APSPVKEKPVMSNIGKSTNHLWNCLIGOSDRQKLLGOKGCVVWITGLSGSGKSTLACA 112

Qy 186 LSRELHCRGHLTYVLDGDNLRHGLNRLDLSFKAEDRAENIRRVGEVAKLFADAGVICIASL 245
Db 113 LSRELHCRGHLTYVLDGDNLRHGLNRLDLSFKAEDRAENIRRVGEVAKLFADAGVICIASL 172

Qy 246 ISPYRRDRDACRALLPHSNFIEVDLPLKICEARDPKGLYKLARTGKIKGFTGIDDPYE 305
Db 173 LSPYRRDRDACRALLPHSNFIEVDLPLKICEARDPKGLYKLARTGKIKGFTGIDDPYE 232

Qy 306 PPINGEIVIKMKDEECPSPKAMAKQVLCYLEENGYLQA 343
Db 233 PPINGEIVIKMKDEECPSPKAMAKQVLCYLEENGYLQA 270

RESULT 3
US-10-425-114-47513
; Sequence 47513, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
```

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 47513
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700097023_FLI.ppep
US-10-425-114-47513

Query Match 55.5%; Score 1013.5; DB 15; Length 253;
Best Local Similarity 79.8%; Pred. No. 1.2e-81;
Matches 210; Conservative 4; Mismatches 20; Indels 29; Gaps 6;

Qy 64 SDAPLALVIHGLTPRSSHSSAGLASDSGRREGGARGA---RT-----HCHRGIGRWV 113
Db 1 SMASLP--VPHTL--PRASPAIVGAA-----RGRAAVRVRTATATGALGGCGGGGMEQ 49

Qy 114 RRRRNGAACEAPHSPVKEKPVMSNIGKSTNHLWNCLIGOSDRQKLLGOKGCVVWITG 173
Db 50 RH-----GEAPHSPVKEKPVMSNIGKSTNHLWNCLIGOSDRQKLLGOKGCVVWITG 101

Qy 174 LSGSGKSTLACALSRELHCRGHLTYVLDGDNLRHGLNRLDLSFKAEDRAENIRRVGEVAKL 233
Db 102 LSGSGKSTLACALSRELHCRGHLTYVLDGDNLRHGLNRLDLSFKAEDRAENIRRVGEVAKL 161

Qy 234 PADAGVICIASLISPYRRDRDACRALLPHSNFIEVDLPLKICEARDPKGLYKLARTGK 293
Db 162 FSDAGVICIASLISPYRRDRDACRALLPHSNFIEVDLPLKICEARDPKGLYKLARTGK 221

Qy 294 IKGFTGIDDPYEPPINGEIVIKM 316
Db 222 IKGFTGIDDPYEPPINGEIRHKL 244

RESULT 4
US-10-739-930-10868
; Sequence 10868, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 10868
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-C7564_1.p
US-10-739-930-10868

Query Match 53.2%; Score 972; DB 16; Length 207;
Best Local Similarity 87.9%; Pred. No. 4.5e-78;
Matches 182; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

Qy 137 MSNIGKSTNHLWNCLIGOSDRQKLLGOKGCVVWITGLSGSGKSTLACALSRELHCRGHL 196
Db 1 MSNIGKSTNHLWNCLIGOSDRQKLLGOKGCVVWITGLSGSGKSTVACALSRELHYRHH 60

Qy 197 TVVLDGDNLRHGLNRLDLSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDAC 256
```

Db 61 TYVLDGDNLRHGLNLSFKAEDRTENIRRVGEVAKLFADAGTICIASLISPYRRDRDAC 120
QY 257 RALLPHSNFIEVFDLPLKICARPKGLYKLRATKIKGFTGIDDPYEPPIGVIM 316
Db 121 RALLPDSRFIEVFDLPLKICARPKGLYKLRATKIKGFTGIDDPYEPPIGVIM 180
QY 317 KDEECPSPKAMAKQVLCYLENGYLQA 343
Db 181 EGGECPSPKAMAKQVLCYLENGYLQA 207

RESULT 5
US-10-767-701-40866
; Sequence 40866, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 40866
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C114933_1.pgp
US-10-767-701-40866

Query Match 51.3%; Score 936; DB 16; Length 188;
Best Local Similarity 94.7%; Pred. No. 6.4e-75;
Matches 178; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
QY 156 SDRQKLGQKGCVMWITGLSGSGKSTLACALSREHLCHRGHTYVLDGDNLRHGLNRLDSF 215
Db 1 SDRQKLGQKGCVMWITGLSGSGKSTLACALSREHLCHRGHTYVLDGDNLRHGLNRLDSF 60
QY 216 KAEDRAENIRRVGEVAKLFADAGVICTIASLISPYRRDRDACRALLPHSNFIEVFDLPLK 275
Db 61 KAEDRAENIRRVGEVAKLFADAGVICTIASLISPYRRDRDACRALLPHSNFIEVFDLPLK 120
QY 276 ICEARDPKGLYKLRATKIKGFTGIDDPYEPPIGVIMKDECPSPKAMAKQVLCYL 335
Db 121 ICEARDPKGLYKLRATKIKGFTGIDDPYEPPIGVIMKDECPSPKAMAKQVLCYL 180
QY 336 BENGYLQA 343
Db 181 BENGYLQA 188

RESULT 6
US-10-425-115-247552
; Sequence 247552, Application US/10425115
; Publication No. US200401272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 247552
; LENGTH: 429
; TYPE: PRT

; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(429)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_157357C.1.pgp
US-10-425-115-247552
Query Match 49.5%; Score 903; DB 16; Length 429;
Best Local Similarity 53.7%; Pred. No. 1.6e-71;
Matches 189; Conservative 33; Mismatches 98; Indels 32; Gaps 4;
QY 7 NOTEPLVTHTOQPP-----SPAGPASQOQROGNTLL-----SPTPTTLAVILV 49
Db 91 SRTPRTAHGRTPPCALVSLVSRLPPPPQARGGGALAILVSTPAAGVFPFPLAVARA 150
QY 50 NPQRAPPVPLGLTPSDAPLPALVTHGLTPRSHSAGLASDSGRREGSGRGARTHCHRG 109
Db 151 SRXRA-----PEPEPRRESRVASASPPPPPETPSGRNXDEGNRRQRLXGRPLVEGAG 202
QY 110 GRWVRRRRNGAAGEAPHSPVKEKPVMSNIGKSTNIIWHNCLIGQSDROKLGKGCVV 169
Db 203 DRSVEBOPEHAGVEGQA-----LXMSSTVPKSSNIFMWDCLVGKTDROKLLNOKGCVV 255
QY 170 WITGLSGSGKSTLACALSREHLCHRGHTYVLDGDNLRHGLNRLDSFKAEDRAENIRRVGE 229
Db 256 WITGLSGSGKSTLACTLGLRELHTRGKLAYVLDGDNLRHGLNKLDFKAEDRAENIRRVGE 315
QY 230 VAKLFADAGVICTIASLISPYRRDRDACRALLPHSNFIEVFDLPLKICEARDPKGLYKLA 289
Db 316 VAKLFADAGVICTIASLISPYRRDRDRESCALLSDSFIEVFLNMSLELCEARDPKGLYKLA 375
QY 290 RTGKIKGFTGIDDPYEPPIGVIMKDECPSPKAMAKQVLCYLENGYL 341
Db 376 RAGKIKGFTGIDDPYEPPIGVIMKDECPSPKAMAKQVLCYLENGYL 427

RESULT 7
US-10-437-963-154262
; Sequence 154262, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 154262
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_5413C.1.pgp
US-10-437-963-154262

Query Match 49.3%; Score 900; DB 16; Length 304;
Best Local Similarity 57.0%; Pred. No. 1.9e-71;
Matches 188; Conservative 27; Mismatches 61; Indels 54; Gaps 4;
QY 23 APGASQOQROGNTLLSPTPTLA--VILVNPQRPVPLPGLTPSDAPLPALVHGLTPRS 80
Db 18 APAAAADGRLAGGVLRPPWGGVAAAGVRLVAAAR-----PMS 53

Qy	14	THTQPPSPA-----PGPASQCORQNTLL-----SPTP--TLAVILNVP	51
Db	10	THASRRRAHPARVRVPLSASRLPPPQRGARGGALAVLSIPAAGVPVPLFLAVARASR	69
Qy	52	QRAPVPLPGLTPSDAPLPAIVIHGLTPRSSHSSAGLASDSGRREGEGRGARTHCRGIGR	111
Db	70	XRAPPEPRVKSRVASAP-----PPPTPSGWNQNEGDRRRKLXGRPLVECTGDR	121
Qy	112	WVRRRRNGAAPGEAPHSPVKEK--PVMSNIKGKTNILWHNCLIGOSDQKLLQGKCVCV	169
Db	122	SVEEQ-----PEHAGXSLVBGKALKMKSSTVPKSSNIFWHDPCVGKTDRQLLKKQGCVCV	175
Qy	170	WITGLSGSKSTLACALSRELHCRCGLTVVLDDGNLRHCLNPEDLSFKAEDRAENIRRVGE	229
Db	176	WITGLSGSKSTLACTLGRELHTRGKLAYVLDDNLRHGLNKOLGFKAEDRAENIRRVGE	235
Qy	230	VAKLFADAGVICIASLSPYRRDRDACRALPHSNFIEVFIDLPUKICEARDPKGLYKLA	289
Db	236	VAKLFADAGLVCIASLSPYRRDRESCHALLSDSSFIEVFLNMSLELCARDPKGLYKLA	295
Qy	290	RTGKIKGFTGTIDDPYPPIINGEIVIMKDCECPSKAMAKQVLCYLENGYL	341
Db	296	RAGKIKGFTGTIDDPYEAPLNCEIETKEVDGVCPPPAEMAGOVVTILEKGFL	347

```

RESULT 11
US-10-829-432-10
; Sequence 10, Application US/10829432
; Publication No. US20040177401A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio
; APPLICANT: Allen, Stephen
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-B
; CURRENT APPLICATION NUMBER: US/10/829,432
; PRIOR FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/720,384A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 246
; TYPE: prt
; ORGANISM: Triticum aestivum
US-10-829-432-10

```

RESULT 12
US-10-425-114-55578

```

; Sequence 55578, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 55578
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700856117_FLI.pep
US-10-425-114-55578

Query Match 46.8%; Score 854; DB 15; Length 215;
Best Local Similarity 74.9%; Pred. No. 1.5e-67;
Matches 155; Conservative 27; Mismatches 25; Indels 0; Gaps 0;

Qy 137 MSNIGKSTNLTWNCILIGSDROKLLIGQKGVVWITGLSGSGKSTLACLSRELHCRGHL 196
| : | | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 9 MATLSNSTNIFWQDCQVGLERQKLNQKGVVWITGLSGSGKSTLACLSRELHSGKGL 68
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 197 TYVLDGDNLRHGLNRDLSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDAC 256
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 69 SYVLDGDNLRHGLNKDLGFKAEADRAENIRRTGEVAKLFADAGLICVASLISPYRRDRDTC 128
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 257 RALLPHSNFTVEPIDPLPKICEARDPKGLYKLUARTGKIKGFTGIDDPYPPPPINGEIVIKM 316
| : | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 129 RAMLPANFTVEPMNPFELCEARDPKGLYKLUARAGIKGFTGIDDPYPPPINCEIEIQ 188
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy 317 KDECPSPKAWAKVLCYLEENGYLQA 343
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 189 ENGDCPTPTLMACQVVTYLENKGFLES 215
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 13
US-10-424-599-205747
; Sequence 205747, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 205747
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_27817C.1.pgp
US-10-424-599-205747

```

RESULT 12

Db 59 MSNVANSTNIMWHDPIQKDRQQLQKQKGVWITGLSGSGKSTIACALSRLHSKGL 118
Qy 197 TVVLGDGNLRHGLNRLDLSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDAC 256
Db 119 SYVLGDGNLRHGLNODLSFRAEDRSNIRRIEVEVAKLFADAGVICITSLISPYQKDRDAC 178
Qy 257 RALLPHSNFIEVFDLPLKICEARDPKGLYKLARTGKIKGFTGIDDPYEPPIGEIVIKM 316
Db 179 RALIPKGFIEVFDVPLHVCEARDPKGLYKLARAGIKSFITGIDDPYEPPIPCSEIVLQO 238
Qy 317 KDEECPSPKAMAKQVLCYLENGYLQA 343
Db 239 KGSCKSPDMAEVEVISYLENGYLRA 265

RESULT 14
US-10-424-599-262462
; Sequence 262462, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 262462
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_79025C.1.pap
US-10-424-599-262462

Query Match 46.7%; Score 853; DB 15; Length 207;
Best Local Similarity 75.4%; Pred. No. 1.7e-67;
Matches 156; Conservative 24; Mismatches 27; Indels 0; Gaps 0;
Qy 137 MSNIGKSTNIMWHDPIQKDRQQLQKQKGVWITGLSGSGKSTIACALSRLHSKGL 196
Db 1 MATLSNSTNIFWQDCQIGRPERQKLNQKGVWITGLSGSGKSTIACALSRLHSKGL 60
Qy 197 TVVLGDGNLRHGLNRLDLSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDAC 256
Db 61 SYVLGDGNLRHGLNODLSFKAEDRTENIRRIEVEVAKLFADAGLICVASLISPYRRDRDTC 120
Qy 257 RALLPHSNFIEVFDLPLKICEARDPKGLYKLARTGKIKGFTGIDDPYEPPIGEIVIKM 316
Db 121 RAMLPDANFIEVFMNMPLELCEARDPKGLYKLARAGIKSFITGIDDPYEPPIPCSEIVLQO 180
Qy 317 KDEECPSPKAMAKQVLCYLENGYLQA 343
Db 181 ENGNCPPTMTMAGQVVTYLENGFLES 207

RESULT 15
US-10-829-432-8
; Sequence 8, Application US/10829432
; Publication No. US20040177401A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio
; APPLICANT: Allen, Stephen
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-B
; CURRENT APPLICATION NUMBER: US/10/829,432
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/720,384A
; PRIOR FILING DATE: 2000-12-21

; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Glycine max
US-10-829-432-8
Query Match 46.7%; Score 852; DB 16; Length 311;
Best Local Similarity 73.9%; Pred. No. 3.6e-67;
Matches 153; Conservative 31; Mismatches 23; Indels 0; Gaps 0;
Qy 137 MSNIGKSTNIMWHDPIQKDRQQLQKQKGVWITGLSGSGKSTIACALSRLHSKGL 196
Db 105 MSNVGNSTNIMWHDPIQKDRQQLQKQKGVWITGLSGSGKSTIACALSRLHSKGL 164
Qy 197 TVVLGDGNLRHGLNRLDLSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPYRRDRDAC 256
Db 165 SYVLGDGNLRHGLNODLSFRAEDRSNIRRIEVEVAKLFADAGVICITSLISPYQKDRDAC 224
Qy 257 RALLPHSNFIEVFDLPLKICEARDPKGLYKLARTGKIKGFTGIDDPYEPPIGEIVIKM 316
Db 225 RALLSKGDFIEVFDVPLHVCEARDPKGLYKLARAGIKSFITGIDDPYEPPIPCSEIVLQO 284
Qy 317 KDEECPSPKAMAKQVLCYLENGYLQA 343
Db 285 KGSCKSPDMAEVEVISYLENGYLRA 311

Search completed: September 9, 2005, 13:07:10
Job time : 113 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 9, 2005, 13:03:11 ; Search time 4068 Seconds
(without alignments)
4085.580 Million cell updates/sec

Title: US-10-829-432-4

Perfect score: 1826

Sequence: 1 RPFHFINQTEPLVTHTOQPP.....PRMAKQVLCYLENGYLQA 343

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10829432/runat_08092005.132223_25075/app_query.fasta_1.519
-DB=GenEmbl -QFMT=fastap -SUFFIX=rgc -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10829432 @CGN_1_1_5600 @runat_08092005.132223_25075 -NCPUL=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*

1: gb_ba:*

2: gb_hgt:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	876	48.0	928	8 BT009518	Triticum
2	874.5	47.9	1434	8 AK071285	Oryza sat
3	833	45.6	1086	8 AF044285	Catharant
4	816	44.7	831	6 CQ805858	Sequence

5	816	44.7	831	6 AX412633	Sequence
6	816	44.7	831	6 AX506015	Sequence
7	816	44.7	831	6 AV132010	Arabidops
8	816	44.7	1068	8 AY085264	Arabidops
9	816	44.7	1077	8 ATU05238	Arabidops
10	816	44.7	1143	8 AY054287	Arabidops
11	816	44.7	1185	8 ATCLAPSK	A.thaliana
12	815.5	44.7	970	8 AF178976	Zea mays
13	778.5	42.6	920	6 BD248389	Gene part
14	778.5	42.6	1243	8 AF462823	Arabidops
15	778.5	42.6	1311	8 AF043351	Arabidops
16	777.5	42.6	882	8 AY097421	Arabidops
17	757.5	41.5	964	8 BT005193	Arabidops
18	757.5	41.5	1143	8 BT003977	Arabidops
19	752.5	41.2	1133	8 AY085031	Arabidops
20	664	36.4	133022	8 ACL18346	Oryza sat
21	633.5	34.7	3489	8 ATU59759	Arabidops
22	633.5	34.7	89840	8 AC004705	Arabidops
23	632.5	34.6	51860	8 AB013390	Arabidops
24	629.5	34.5	112067	8 ATCHRIV92	Arabidops
25	629.5	34.5	118267	8 ATT5J17	Arabidops
26	599	32.8	110000	1 AF006618_26	Continuation (27 o
27	595.5	32.6	1017	8 AF069952	Enteromor
28	592.5	32.4	302325	1 AE017236	AE017236 Mycobacte
29	587.5	32.2	101410	8 ATAC009540	Arabidops
30	577	31.6	11131	1 AS007523	Arabidops
31	572.5	31.4	58756	2 AP006484_4	Clostridi
32	565.5	31.0	53533	1 BSY09476	Continuation (5 of
33	565.5	31.0	198743	1 BSUB00006	Y09476 Bacillus su
34	563	30.8	2372	3 AK173431	Z99109 Bacillus su
35	563	30.8	2378	3 AB036852	AK173431 Ciona int
36	560	30.7	192657	1 AP005377	AB036852 Ciona int
37	556.5	30.5	110000	1 AE017282_23	AP005377 Thermosyn
38	552	30.2	153950	8 AC104429	Continuation (24 o
39	551	30.2	2380	3 AK173392	AC104429 Oryza sat
40	549.5	30.1	300150	1 AP004598	AK173392 Ciona int
41	549	30.1	147205	8 AC146702	AP004598 Oceanobac
42	547	30.0	1045	8 AV325892	AC146702 Genomic s
43	542	29.7	30275	1 AE016751	AV325892 Porphyra
44	538.5	29.5	302050	1 AL935256	AE016751 Staphyloc
45	537.5	29.4	71349	8 AP006662	AL935256 Lactobaci
					AP006662 Lotus cor

ALIGNMENTS

RESULT 1	BT009518	Triticum aestivum clone wr1.pk0101.e2:fis, full insert mRNA	928 bp	linear	PLN 20-JUN-2003
LOCUS	BT009518	sequence.			
DEFINITION	BT009518	sequence.			
ACCESSION	BT009518				
VERSION	BT009518.1	GI:32129069			
KEYWORDS	FLI CDNA.				
SOURCE	Triticum aestivum (bread wheat)				
ORGANISM	Triticum aestivum				
REFERENCE	1 (bases 1 to 928)				
AUTHORS	Tingey,S.V., Wolters,P., Powell,W., Dolan,M., Miao,G.-H., Caraher,N.R., Hanafey,M.K. and Hainey,C.F.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA				
FEATURES	Location/Qualifiers				
source	1..928				
	/organism="Triticum aestivum"				
	/mol_type="mRNA"				
	/db_xref="taxon:4565"				
	/clone="wr1.pk0101.e2:fis"				
ORIGIN					

[illegible]

Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Komno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishii, K., Nomura, K., Numasaki, R., Ohno, M., Otsu, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers
 1. .1434
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J023086D14"

FEATURES

source

Alignment Scores:

Pred. No.: 2,23e-33 Length: 1434
 Score: 874.50 Matches: 191
 Percent Similarity: 63.74% Conservative: 34
 Best Local Similarity: 54.11% Mismatches: 68
 Query Match: 47.89% Indels: 62
 DB: 8 Gaps: 7

US-10-829-432-4 (1-343) x AK071285 (1-1434)

QY 49 ValAsnProGlnArgAlaProProValLeuProGlyLeuThrProSerAspAlaProLeu 68
 DB 6 GTTAAACCCATTCGGGGGCTCCATCGATGGAGGCTCTCTCCCGTTTACACACACCA 65
 QY 69 ProAlaLeuValIleHis--GlyLeuThr-----ProArgSerSerHisSer 84
 DB 66 CCCCCTGCTCTCTCACCGCGGCGATCACCGGCGCCCTCATCTCTCTCTCTCTCTCC 125
 QY 85 Ala-----GlyLeuAlaSerAspSerGlyArgGluGlyGly 98
 DB 126 GCGAGACCGCGTGCACCTGCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 185
 QY 99 ArgGlyAlaArgThr-----GlyLeuAlaSerAspSerGlyArgGluGlyGly 103
 DB 186 GAGCCCGGCAATATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 245
 QY 104 -----HisCysHisArgGlyIleGlyArg----- 111
 DB 246 GCGCGGAGGATAACGCTGTCAGTCACGGAG-GGCGCGTGGTGGTGGTGGTGGTGGTGG 304
 QY 112 -----Trp----- 112
 DB 305 CGCCTCTCTGGATCGCGCGCGGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 364
 QY 113 ---ValArgArgArgArgArgAlaAlaProGlyGluAlaProHisSerProVal 131
 DB 365 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 423
 QY 132 LysGluLysProValMet-----SerAsnIleGlyLysSerThrAsnIleLeuTrp 148
 DB 424 GCGGAGAACAAAGTTTGCATATGTCATCAATGTGCGGAGGCGTCCAAATCTCTGG 483
 QY 149 HisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuLeuGlyGlnLysGlyCysVal 168
 DB 484 CATGATTGTCAGTTGGCAGCTGATCGCAGAGCTACTGAACGAGAAAGTTGCGTT 543
 QY 169 ValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAlaCysAlaLeuSerArg 188
 DB 544 GTTGTGATCACAGGACTAGTGGTTCAGGTAAGTACCTGGCATGCACATAGATCGA 603
 QY 189 GluLeuHisCysArgGlyHisLeuThrThrValLeuAspGlyAspAsnLeuArgHisGly 208
 DB 604 GAGCTCCATACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 663
 QY 209 LeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIleArgArgValGly 228

DB 664 TTGAACAGGATCTTGGCTTTAAAGCGGAGACCGCTGCTGCTGCTGCTGCTGCTGCTGCT 723
 QY 229 GluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSerPro 248
 DB 724 GAGGTAGCAAGAGCTATTTCGAGATCGAGCTGATGATGATGATGATGATGATGATGAT 783
 QY 249 TyrArgArgAspArgAspAlaCysArgAlaLeuLeuProHisSerAsnPheIleGluVal 268
 DB 784 TATAGGAGAGACCGTGAAGTCTTTCGTCATATTGTCAGATGGTAGCTTTATTCAGTT 843
 QY 269 PheIleAspLeuProLeuLysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeu 288
 DB 844 TTCTTGAACATGCTTGGATTTGTGAGTCAAGGATCTCTAAGGGCTTTTATAAGCTT 903
 QY 289 AlaArgThrGlyLysIleLysGlyPheThrGlyLysAspAspProTyrGlyProPheIle 308
 DB 904 GCTCGTCAGGAGAAATAAGGGTTTTACTGGAATTGATGACCTTTATGAATCACCATT 963
 QY 309 AsnGlyGluIleValIleLysMetLysAspGluGluCysProSerProLysAlaMetAla 328
 DB 964 AACTCTGAGATTGATCACAAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1023
 QY 329 LysGlnValLeuCysTyrLeuGluGluAsnGlyTyrLeu 341
 DB 1024 GGCACAAAGTAGTCACTTATCTAGAGGAGAAAGGGTTTCTG 1062

RESULT 3
 AF044285 1086 bp mRNA linear JUN 11-AUG-1998
 LOCUS Catharanthus roseus adenosine-5'-phosphosulfate-kinase (CRakn)
 DEFINITION mRNA, complete cds.
 ACCESSION AF044285
 VERSION AF044285.1 GI:2832299
 KEYWORDS Catharanthus roseus (Madagascar periwinkle)
 ORGANISM Catharanthus roseus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Gentianales; Apocynaceae; Rauvolfioideae;
 Vincet; Catharanthus.
 REFERENCE 1 (bases 1 to 1086)
 AUTHORS Arz, H.E., Gisselmann, G., Schiffmann, S. and Schwenn, J.D.
 TITLE A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana
 JOURNAL Biochim. Biophys. Acta 1218 (3), 447-452 (1994)
 MEDLINE 94325358
 PUBMED 8049272
 REFERENCE 2 (bases 1 to 1086)
 AUTHORS Schiffmann, S. and Schwenn, J.D.
 TITLE Isolation of cDNA clones encoding
 adenosine-5'-phosphosulfate-kinase (EC2.7.1.25) from Catharanthus
 roseus (Accession No. AF044285) and an isoform (akn2) from
 Arabidopsis (Accession No. AF043351) (PGR98-116)
 JOURNAL Plant Physiol. 117 (3), 1125 (1998)
 REFERENCE 3 (bases 1 to 1086)
 AUTHORS Schiffmann, S. and Schwenn, J.D.
 TITLE Direct Submission
 JOURNAL Submitted (23-JAN-1998) Biochemie der Pflanzen, Ruhr-Universität
 Bochum, Universitätsstrasse 150, Bochum 44780, Germany
 FEATURES
 source
 1. .1086
 /organism="Catharanthus roseus"
 /mol_type="mRNA"
 /db_xref="taxon:4058"
 /sub_clone="KSCRakn"
 /note="heterotrophic cell culture grown under sulfate
 deficiency"
 1. .1086
 /gene="CRakn"
 39. .977
 /gene="CRakn"
 /EC_number="2.7.1.25"
 /function="phosphorylates 3'-OH group of
 adenosine-5'-phosphosulfate"


```

Qy 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
Db 757 GGAGGAACCTTCTCTATCGAAATGCGGAAAGGTCGTGGATACCTTAGATAACAAGGCT 816

Qy 340 TyrLeuGlnAla 343
Db 817 TATCTTCAAGCA 828

RESULT 5
AX412633
LOCUS AX412633 831 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 397 from Patent WO0222675.
ACCESSION AX412633
VERSION AX412633.1 GI:21445091
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS Glazebrook, J., Wang, X., Dangl, J. L., Eulgem, T. and Zhu, T.
TITLE Plant genes, the expression of which are altered by pathogen
infection
JOURNAL Patent: WO 0222675-A 397 21-MAR-2002;
Syngenta Participations AG (CH); UNIVERSITY OF NORTH CAROLINA AT
CHAPEL HILL (US); Glazebrook, Jan (US); Wang, Xun (US); Dangl,
Jeffrey L. (US); Eulgem, Thomas (US)
FEATURES
source
1..831
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores: 7,94e-31 Length: 831
Pred. No.: 816.00 Matches: 159
Score: 77.87% Conservative: 31
Percent Similarity: 65.16% Mismatches: 44
Best Local Similarity: 44.69% Indels: 10
Query Match: 6 Gaps: 3
DB:

US-10-829-432-4 (1-343) x AX412633 (1-831)

Qy 100 GlyAlaArgThrHisCysHisArgGlylleGlyArgTrpValArgArgArgArgAsn 119
Db 127 GGATCTCAAACTCTGAGTCATACAAAATGGAATGATCTATTCTCTGAGGTAAATCCATTAA 186

Qy 120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139
Db 187 GGT-----CACACGGGACAAAAGCAAGGACCT---TTGTCTACG 222

Qy 140 lleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
Db 223 GTCGGAACCTCGACAAATATAAAGTGGCATGAATGTTCTCTGTGAGAAAGTTGATAGACAG 282

Qy 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
Db 283 AGATTGCTTGATCAGAAAAGGATGTGTGATTTGGGTCCCGGCTTTAGTGGTTTCAGGGAAG 342

Qy 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpVal 199
Db 343 AGTACTTTGGCTTGCTTGATGATCAGATGTTGTATCAAAAGGGGAAGCTTTGTTATAT 402

Qy 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
Db 403 CTTGATGGTGATAATGTTAGGCATGCTTTAAACCGTGATCTTTAGCTTTAAAGCTCAGGAT 462

Qy 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
Db 463 CGTCGAGAGAATATTTCGTAGAGTTGGAGAGGTTGCTTAAGCTTTTTCGCGAGTCTGGAATA 522

```

```

Qy 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
Db 523 ATCTGCATTGCGAGTTTGATATCTCTTATAGACAGATAGGACGCTTGTCGAAGTTTG 582

Qy 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
Db 583 CTCCCGCAGGAGATTTTGTGAGGTGTTCAATGATGTACCGCTTAGTGTTCGAGGCG 642

Qy 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
Db 643 AGGATCCAAAGGCTCTTACAGCTTCTGTCGAGGAAGATCAAGGTTTTCACGGG 702

Qy 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMethLysAspGlu 319
Db 703 ATCGATGACCTTACGAGCCACCATTTGAATGCGAG-----ATTCTCTAGACGTGAA 756

Qy 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
Db 757 GGAGGAACCTTCTCTATCGAAATGCGGAAAGGTCGTGGATACCTTAGATAACAAGGCT 816

Qy 340 TyrLeuGlnAla 343
Db 817 TATCTTCAAGCA 828

RESULT 6
AX506015
LOCUS AX506015 831 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 710 from Patent WO0216655.
ACCESSION AX506015
VERSION AX506015.1 GI:23387252
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS Harper, J. F., Kreps, J., Wang, X. and Zhu, T.
TITLE Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 710 28-FEB-2002;
The Scripps Research Institute (US); Syngenta Participations AG
(CH)
FEATURES
source
1..831
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores: 7,94e-31 Length: 831
Pred. No.: 816.00 Matches: 159
Score: 77.87% Conservative: 31
Percent Similarity: 65.16% Mismatches: 44
Best Local Similarity: 44.69% Indels: 10
Query Match: 6 Gaps: 3
DB:

US-10-829-432-4 (1-343) x AX506015 (1-831)

Qy 100 GlyAlaArgThrHisCysHisArgGlylleGlyArgTrpValArgArgArgArgAsn 119
Db 127 GGATCTCAAACTCTGAGTCATACAAAATGGAATGATCTATTCTCTGAGGTAAATCCATTAA 186

Qy 120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139
Db 187 GGT-----CACACGGGACAAAAGCAAGGACCT---TTGTCTACG 222

Qy 140 lleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
Db 223 GTCGGAACCTCGACAAATATAAAGTGGCATGAATGTTCTGTGAGAAAGTTGATAGACAG 282

Qy 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
Db 283 AGATTGCTTGATCAGAAAAGGATGTGTGATTTGGGTCCCGGCTTTAGTGGTTTCAGGGAAG 342

Qy 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpVal 199
Db 343 AGTACTTTGGCTTGCTTGATGATCAGATGTTGTATCAAAAGGGGAAGCTTTGTTATAT 402

Qy 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
Db 403 CTTGATGGTGATAATGTTAGGCATGCTTTAAACCGTGATCTTTAGCTTTAAAGCTCAGGAT 462

Qy 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
Db 463 CGTCGAGAGAATATTTCGTAGAGTTGGAGAGGTTGCTTAAGCTTTTTCGCGAGTCTGGAATA 522

```

```

Db      283 AGATTGCTTGATCAGAAAGGATGTGTGATTTGGGTACCGGTCTTAGTGGTTACGGAAG 342
Qy      180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
Db      343 AGTACTTTGGCTGTGCTTGAATCAGATGTTGTATCAAAAGGGGAAGCTTTGTTATATT 402
Qy      200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLeuAlaGluAsp 219
Db      403 CTTGATGGTGATAAGTTAGGATGGCTTAACCGTGATCTTAGCTTTAAGCTGAGGAT 462
Qy      220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
Db      463 CGTCGACAGAAATATTCGTAGAGTTCGAGAGGTTGCTAAGCTTTTTCGCGATGCTGAATA 522
Qy      240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgGluAlaCysArgAlaLeu 259
Db      523 AFTCTGATTCGAGGTTGATCTCTCTATAGAACAGATAGGAGCGCTTGTCGAAGTTTG 582
Qy      260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
Db      583 CTCCTCCCGGAGAGATTGTTGAGGTGTTTCATGGATGTACCGCTTAGTGTTCGAGGG 642
Qy      280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
Db      643 AGGGATCCAAAGGCTCTTTACAAAGCTTGTCTGTCGAGGAAGATCAAAAGGTTTACCGG 702
Qy      300 IleAspAspProTyrGluProProlIleAsnGlyGluIleValIleLysMetLysAspGlu 319
Db      703 ATCGATGACCCCTTAGAGCCCATGTAACCTCGAG-----ATTCTCTAGGACGTGAA 756
Qy      320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysValTyrLeuGluGluAsnGly 339
Db      757 GGAGNACTTCTCCATCGAATGCGGAAGGTGCTCGGATCTTAGATAACACAGG 816
Qy      340 TyrLeuGlnAla 343
Db      817 TATCTTCAAGCA 828

RESULT 7
AY132010
LOCUS      AY132010
DEFINITION Arabidopsis thaliana At2g14750/F26C24.11 mRNA, complete cds.
ACCESSION AY132010
VERSION    AY132010.1 GI:22135772
KEYWORDS   FLI CDNA.
SOURCE     Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 831)
AUTHORS    Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Banh,J., Bowser,L.,
            Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
            Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
            and Ecker,J.R.
            Arabidopsis ORF clones
            Unpublished
            2 (bases 1 to 831)
AUTHORS    Kim,C.J., Chen,H., Cheuk,R., Shinn,P., Banh,J., Bowser,L.,
            Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
            Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
            Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
            Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
            Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
            Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
            and Ecker,J.R.
            Direct Submission
            Submitted (26-JUN-2002) Salk Institute Genomic Analysis Laboratory

```

```

(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PCEC (SSP) Consortium members constructed and
sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim,C.J.,
Chen,H., Cheuk,R., Shinn,P., Banh,J., Bowser,L., Chang,E.,
Dale,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lam,B.,
Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J.,
Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,
Yamada,K., Yamamura,Y., Yu,G., Yu,S., Davis,R.W., Theologis,A.,
and Ecker,J.R.

Kim,C.J. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES             Location/Qualifiers
     source            1..831
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="2"
                     /clone="U13881"
                     /ecotype="Columbia"
                     /note="This clone is in pUNI 51"
     CDS               1..831
                     /note="putative adenosine phosphosulfate kinase"
                     /codon_start=1
                     /product="At2g14750/F26C24.11"
                     /protein_id="AA091043.1"
                     /db_xref="GI:22135773"
                     /translation="MIAAGAKSLGLSMASPKGIFDSNSMSNSRSVVVRACVSMDS
                     QTLSHNKGSIPEVKSINGHTQKQGLSTVGNSTKWHSESVKVDRLDLQKGC
                     VIVTGLSGSKSLACALNMLYQKGLCYILDGNVHGLNRLDLSFKADRAENIR
                     RVGEVAKLFADAGIICIASLISPYTRDRDACRSLLPEDGFVEFMDVPLSVCEARDPK
                     GLYKLARAGKIKGFTGIDDDPEPPLNCEISLREGGTSPIEMAEEKVWGLDNKGYLQA
                     "

ORIGIN
Alignment Scores:
Pred. No.:          7,94e-31          Length:          831
Score:              816.00            Matches:         159
Percent Similarity: 77.87%            Conservative:    31
Best Local Similarity: 65.16%          Mismatches:     44
Query Match:        44.69%            Indels:         10
DB:                  8                 Gaps:           3

US-10-829-432-4 (1-343) x AY132010 (1-831)
Qy      100 GlyAlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgArgAsn 119
Db      127 GGAATCTCAACTCTGAGTCATACAAAATGATCTATTCTCTGAGGTTAAATTCATTAC 186
Qy      120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139
Db      187 GGT-----CACCGGGACAAAGCAAGGACCT---TTGTCTACG 222
Qy      140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
Db      223 GTCGGAACTCGACAAATATAAAGTGGCATGAATGTTCTGTGTGAGAAAGTTGATAGACAG 282
Qy      160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
Db      283 AGATTGCTTGTATCAGAAAGGATGTGTGATTTGGGTCACCGCTCTTAGTGGTTACGGAAG 342
Qy      180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
Db      343 AGTACTTTGGCTTTGCTTGAATCAGATGTTGTATCAAAAGGGGAAGCTTTGTTATATT 402

```

sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the Ws or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Geneset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES

Location/Qualifiers
 1..1068
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="14216"
 94..924
 /codon_start=1
 /product="putative adenosine phosphosulfate kinase"
 /protein_id="AA062496.1"
 /db_xref="GI:21553403"
 /translation="MIAAGAKSLQGLSLMASPKGIFDSNMSNSRSVVVVRACVSDGS QTLHNKNGSIPEVKSINGHTGQKQGLPSTVGNSTNFKWHECSVEKVDRLDQKGC VIWVTLGSGSKSLACALNOMLYQKGLCYILDNDVNRHGLNDRDLSFKPADRAENIR RVGSAKLFPADAGIICIASLISPYRTDRCRSLIPSGDFVEFMDVPLSCEARDPK GLYKLARAGKIKGFTGIDDDPYEPPLNCEISLREGGTSPIEMAEKVGVLYLQKYLQA"

ORIGIN

Alignment Scores:
 Pred. No.: 1,02e-30 Length: 1068
 Score: 816.00 Matches: 159
 Percent Similarity: 77.87% Conservative: 31
 Best Local Similarity: 65.16% Mismatches: 44
 Query Match: 44.69% Indels: 10
 DB: 8 Gaps: 3

US-10-829-432-4 (1-343) :x AY085264 (1-1068)

Qy 100 GlyAlaArgThrHisCysHisArgGlyLeuGlyArgTrpValArgArgArgAen 119
 Db 220 GGATCTCAAACTCTGAGTCATAACAAAATGATCTATCTCTGAGGTTAAATCCATTAC 279
 Qy 120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAen 139
 Db 280 GGT-----CACACGGGACAAAAGCAAGGACCT---TTGTCTACG 315
 Qy 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
 Db 316 GTCGAAACTCGACAAATATAAGTGGCATGAATGTTCTGTGTGAGAAAGTTGATAGACAG 375
 Qy 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
 Db 376 AGATTGCTTGATCAGAAAGGATGTGATTGGGTCCACCGCTCTTAGTGGTTCAGGGAAG 435
 Qy 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpVal 199
 Db 436 AGTACTTGTGGCTTGTGTTGATCAGATGTGTATCAAAAGGGAAGCTTTGTATATT 495
 Qy 200 LeuAspGlyAspAenLeuArgHisGlyLeuAsnArgAspLeuSerPheLeuAlaGluAsp 219
 Db 496 CTTGATGGTGATAATGTTAGGCATGGCTTAAACCGTGATCTTAGCTTTAAAGCTGAGGAT 555
 Qy 220 ArgAlaGluAenIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
 Db 556 CGTCAGAGATATTCGTAGAGTTGGAGAGTTTCTTAAGCTTTTTCGGATGCTGGAATA 615
 Qy 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
 Db 616 ATCTGCTTGGAGTTGATATCTCTTATAGAACAGATAGGACGCTTGTGCAAGTTTG 675
 Qy 260 LeuProHisSerAenPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
 Db 676 CTCCTCCGAGGAGATTTTGTGAGGTGTTTCATGATGTACCGCTTAGTGTTCGAGGCG 735
 Qy 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299

Qy 200 LeuAspGlyAspAenLeuArgHisGlyLeuAenArgAspLeuSerPheLysAlaGluAsp 219
 Db 403 CTTGATGGTGATAATGTTAGGCATGGCTTAAACCGTGATCTTAGCTTTAAAGCTGAGGAT 462
 Qy 220 ArgAlaGluAenIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
 Db 463 CGTCAGAGATATTCGTAGAGTTGGAGAGTTGCTTAAGCTTTTTCGGATGCTGGAATA 522
 Qy 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
 Db 523 ATCTGATTCGAGTTGATATCTCTTATAGAACAGATAGGACGCTTGTGCAAGTTTG 582
 Qy 260 LeuProHisSerAenPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
 Db 583 CTCCTCCGAGGAGATTTTGTGAGGTGTTTCATGAGTGTACCGCTTAGTGTTCGAGGCG 642
 Qy 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
 Db 643 AGGATTCCAAAGGCTCTTTACAGCTTGTCTGTCAGGAAAGATCAAAAGGTTTACCGGG 702
 Qy 300 IleAspAspProTyrGluProIleAenGlyGluIleValIleLysMetLysAspGlu 319
 Db 703 ATCGATGACCTTACGAGCCCATGTGAACCTGCGAG-----ATTCTCTAGGACGTGAA 756
 Qy 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAenGly 339
 Db 757 GGAGGACCTTCTCTATCGAATGCGGGAAGGTCGTCGGATATCTAGATAACAGGCT 816
 Qy 340 TyrLeuGlnAla 343
 Db 817 TATCTTCAAGCA 828

RESULT 8

AY085264 1068 bp mRNA linear PLN 14-APR-2003
 Arabidopsis thaliana clone 14216 mRNA, complete sequence.
 AY085264
 VERSION AY085264.1 GI:21403974
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1068)
 Haas,B.J., Volkovskiy,N., Town,C.D., Trukhan,M., Alexandrov,N., Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
 Full-length messenger RNA sequences greatly improve genome annotation
 Genome Biol. 3 (6), RESEARCH0029 (2002)
 22088475
 2 (bases 1 to 1068)
 Brover,V., Trukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.
 Full-length cDNA from Arabidopsis thaliana
 Unpublished
 3 (bases 1 to 1068)
 Brover,V., Trukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and Feldmann,K.
 Direct Submission
 Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road, Malibu, CA 90265, USA
 This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unspliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A

```
Db 736 AGGGATCCAAAGGGCTTTACAAAGCTTGTCTGTCAGGAAGATCAAAAGGTTTACCGGG 795
Qy 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
Db 796 ATCGATGACCCCTTAGCGGCCACCTTAGAATCGAG-----ATTCTCTAGGACGTGAA 849
Qy 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
Db 850 GGAGGAACCTTCTCTATCGAATGCGGAAGAGTCTGCGGNATCTTAGATAACAAAGGT 909
Qy 340 TyrLeuGlnAla 343
Db 910 TATCTTCAAGCA 921

RESULT 9
ATU05238
LOCUS Arabidopsis thaliana APS kinase mRNA linear PLN 23-JUN-1998
DEFINITION Arabidopsis thaliana APS kinase mRNA, complete cds.
ACCESSION U05238
VERSION U05238.1 GI:450234
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1077)
Jain,A. and Leustek,T.
A cDNA clone for 5'-adenylylphosphosulfate kinase from Arabidopsis
thaliana
Plant Physiol. 105 (2), 771-772 (1994)
94345022
8066145
2 (bases 1 to 1077)
Lee,S. and Leustek,T.
APS kinase from Arabidopsis thaliana: genomic organization,
expression, and kinetic analysis of the recombinant enzyme
Biochem. Biophys. Res. Commun. 247 (1), 171-175 (1998)
98300303
9636674
3 (bases 1 to 1077)
Leustek,T.
Direct Submission
Submitted (18-JAN-1994) Thomas Leustek, Center for Agricultural
Molec. Biology, Rutgers University, Cook College, College Farm
Road, New Brunswick, NJ 08903, USA
Location/Qualifiers
1. .1077
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="ATK"
/tissue_type="whole plant"
33. .863
/EC_number="2.7.1.25"
/function="phosphorylates 5'-adenylylsulfate on
3'-hydroxyl group"
/note="similar to yeast APS kinase (MET14): SwissProt
Accession Number Q02196; similar to Escherichia coli APS
kinase (cysC): SwissProt Accession Number P23846;
adenosine 5'-phosphosulfate kinase"
/codon_start=1
/evidence=experimental
/product="APS kinase"
/protein_id="AAC50035.1"
/db_xref="GI:450235"
/translation="MIAAGAKSLGLSMASPKGIFDSNSMSRSVVVVRACVSMGDS
OTLSHNKNGSIPVKSIINGHTGQKGPLSTVGNSTNIKWHECSVKVDORQLDQKGC
VTWVTLGSGSGKSTLACALNOMLYOKGLCYILDGDNVRHGLNRDLSPKABDRAENIR
RVGEVAKLFADAGLICIASLISPYTRDRDACSLLPEDGFVEVENDVPLSCVCEARDPK
GLYKLARAGKNGFTIDDPYEPPLNCEISLUREGGSFPIEWAERKVVGLNDKGLQA
"
```

```
transit_peptide 33. .143
/note="potential chloroplast transit peptide with a
protease cleavage site"
mat_peptide 144..860
/product="APS kinase"
misc_feature 354..377
/note="ATP/GTP binding site motif"

ORIGIN
Alignment Scores: 1.03e-30 Length: 1077
Pred. No.: 100 GlyAlaargThrHisCysHisArgGlyIleGlyArgTyrValArgArgArgAsn 119
Score: 816.00 Matches: 159
Percent Similarity: 77.87% Conservative: 31
Best Local Similarity: 65.16% Mismatches: 44
Query Match: 44.69% Indels: 10
DB: 8 Gaps: 3

US-10-829-432-4 (1-343) x ATU05238 (1-1077)
Qy 100 GlyAlaargThrHisCysHisArgGlyIleGlyArgTyrValArgArgArgAsn 119
Db 159 GGATCTCAAACTCTGAGTCATAACAAAATGGATCTATCTCTGAGGTTAAATCCATTAAC 218
Qy 120 GlyAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139
Db 219 GGT-----CACACGGGACAAAAGCAAGGACCT---TTGTCTACG 254
Qy 140 IleGlyLysSerThrAsnIleLeuTyrHisAsnCysLeuIleGlyGlnSerAspArgGln 159
Db 255 GTCGGAACCTCGACAAATATAAAGTGGCATGAATGTTCTGTGTGAGAAAGTTGATAGACAG 314
Qy 160 LysLeuLeuGlyGlnLysGlyCysValValTyrPileThrGlyLeuSerGlySerGlyLys 179
Db 315 AGATTGCTTGATCAGAAAGGATGTGTGATTGGCTACCGGCTCTTAGTGGTTCCAGGAAG 374
Qy 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
Db 375 AGTACTTTGGCTTGTGCTTTGAATCAGATGTGTATCAAAAAGGGGAAAGCTTTTATATT 434
Qy 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
Db 435 CTTGATGGTGATAATGTAGGCATGGCTTAAACCGTGATCTTAGCTTTAAAGCTGAGGAT 494
Qy 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
Db 495 CGTCAGAGAATATTCTGTAGAGTTGGAGAGGTTCTTAAGCTTTTTCGGATGCTGGAATA 554
Qy 240 IleCysIleAlaSerLeuIleSerProTyrThrArgAspArgAspAlaCysArgAlaLeu 259
Db 555 ATCTGCATTGCGAGTTTGATATCTCTTATAGAACAGATAGGACGCTTGTGCAAGTTTG 614
Qy 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
Db 615 CTCCCGAGGAGAGATTGTGAGGTGTTCAAGGATGACCGCTTAGTGTGTTGCGAGCG 674
Qy 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
Db 675 AGGATCCAAAGGCTTTTACAAGCTTGCTCGTCAGGAAAGATCAAAAGGTTTACCGGG 734
Qy 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
Db 735 ATCGATGACCCCTTAGCGGCCACCTTAGAATCGAG-----ATTCTCTAGGACGTGAA 788
Qy 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
Db 789 GGAGGAACCTTCTCTATCGAATGCGGAAGAGTCTGCGGNATCTTAGATAACAAAGGT 848
Qy 340 TyrLeuGlnAla 343
Db 849 TATCTTCAAGCA 860

RESULT 10
AT054287
```


LOCUS AY054287 1143 bp mRNA linear PLN 30-SEP-2001
 DEFINITION Arabidopsis thaliana At2g14750/F26C24.11 mRNA, complete cds.
 ACCESSION AY054287
 VERSION
 KEYWORDS
 SOURCE FLI CDNA.
 ORGANISM Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 1 (bases 1 to 1143)
 Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
 Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
 Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
 Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
 Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
 Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
 Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
 Davis,R.W., Theologis,A. and Ecker,J.R.
 Arabidopsis cDNA clones
 2 (bases 1 to 1143)
 Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
 Bowser,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
 Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
 Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
 Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
 Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
 Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
 Davis,R.W., Theologis,A. and Ecker,J.R.
 Direct Submission
 Submitted (22-AUG-2001) Salk Institute Genomic Analysis Laboratory
 (SIGNAL), Plant Biology Laboratory, The Salk Institute for
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
 USA
 RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
 Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
 Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
 Hayashizaki,Y. and Shinozaki,K.
 The Salk, Stanford, PGEC (SSP) Consortium members carried out the
 sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H.,
 Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J., Bowser,L.,
 Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G.,
 Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M.,
 Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A.,
 Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S.,
 Davis,R.W., Theologis,A., and Ecker,J.R.
 Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
 this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
 contributed equally to this work as PIs.
 Location/Qualifiers
 1..1143
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosomes="2"
 /clone="RAFL09-19-L13 (R13881)"
 /ecotype="Columbia"
 1..92
 93..923
 /note="putative adenosine phosphosulfate kinase"
 /codon_start=1
 /product="At2g14750/F26C24.11"
 /protein_id="AAI06946.1"
 /db_xref="GI:15810038"
 /translation="MTAGAKSLGLSWASPKGIFDSSMSNRSVVVVRACVMDGS
 QTLHNKGSIFEVKTSINGHTCKQKPLSTVGNSTNWKHSCSEKVRQRLQKGC
 VLVWTLGSGSGKSLACLNOMLYOKGLCYLLDGNVRRHGLNRLDSPADRANR
 RVGEVAFADAGIICIASLSPYTRDACLRLPEGDFVEFVDFVFLVCEARDPK
 GYLKLARAKIGFTGIDDPVEPLNCEISLREGGTSPIEMAEKVVVGLDNKGYLQA
 5'UTR
 CDS

3'UTR
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.09e-30 Length: 1143
 Score: 816.00 Matches: 159
 Percent Similarity: 77.87% Conservative: 31
 Best Local Similarity: 65.16% Mismatches: 44
 Query Match: 44.69% Indels: 10
 Gaps: 3
 DB:
 US-10-829-432-4 (1-343) x AY054287 (1-1143)
 QY 100 GYAlaAAGThrHisCysHisArgGlyLeGlyArgTdpValArgArgArgArgaen 119
 Db 219 GGAUCTCAAACTCGAGTCATAACAAAATGATCTATTCTCGAGGTAAATCCATTAAAC 278
 QY 120 GYAlaAAlaProGlyGluAlaProHisSerProValHisGlySerProValMetSerAen 139
 Db 279 GGT-----CACACGGGACAAAAGGAGGACCT--TTGCTACG 314
 QY 140 lLeGlyLysSerThrAsnIleLeuTrpHisCysLeuIleGlyGlnSerAspArgGln 159
 Db 315 GTCCGAAACTCGACAAATATAAAGTGGCATGAATGTTCTGTGAGAAAAGTTGATAGACAG 374
 QY 160 LysLeuLeuGlyGlnLysGlyCysValTrpIleThrGlyLeuSerGlySerGlyLys 179
 Db 375 AGATTGCTTGATCAGAAAGGATGTGATTTGGGTGTCACCGGTCTTAGTGGTTCAGGGAAG 434
 QY 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
 Db 435 AGTACTTTGGCTTGTCTTGAATCAGATGTTGTATCAAAAGGGGAGGCTTTGTTATATT 494
 QY 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelyAlaGluAep 219
 Db 495 CTTGATGCTGATAAATTTAGGCGATGGCTTAAACCGTGATCTTAGCTTTAAAGCTGAGGAT 554
 QY 220 ArgAlaGluAenIleAArgValGlyValAlaLysLeuPheAlaAspAlaGlyVal 239
 Db 555 CGTCAGAGAAATATTCTGAGAGTTGGAGAGTGTCTAAGCTTTTTCGGGATGCTGGAATA 614
 QY 240 lLeCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
 Db 615 ATCTGCATTGCGAGTTGATATCTCTTATAGAACAGATAGGAGCGCTTGTCGAGTTTG 674
 QY 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
 Db 675 CTCCCGGAGGAGATTTTGTGAGGTTCATGATGTACCGCTTAGTGTGTCGAGGCG 734
 QY 280 ArgAspProLysGlyLeuTyrLysLeuAlaAargThrGlyLysIleLysGlyPheThrGly 299
 Db 735 AGGATCCAAAGGCTTTTCAAGCTCTCGTCAGGAAAGATCAAAAGGTTTATCCCGGG 794
 QY 300 lLeAspAspProTyrGluProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
 Db 795 ATCGATGACCTTACGAGCCACCATGATGCGAG-----ATTCTCTAGACGTGAA 848
 QY 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAenGly 339
 Db 849 GGAGGAATCTTCTCTATCGAAATGCGGAAAGGTCGTGCGGATACCTTAGATAACAAGGCT 908
 QY 340 TyrLeuGlnAla 343
 Db 909 TATCTTCAAGCA 920
 RESULT 11
 ATCLAPS
 LOCUS ATCLAPS 1185 bp mRNA linear PLN 05-FEB-1998
 DEFINITION A.thaliana (L.Heynh.) chloroplast mRNA for recombinant APS-kinase.
 ACCESSION X75782
 VERSION X75782.1 GI:414736
 KEYWORDS APS-kinase; ATP:adenylylsulfate-3'-phosphotransferase.

SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE	1
AUTHORS	Arz, H.E., Gisselmann, G., Schiffmann, S. and Schwenn, J.D.
TITLE	A CDNA for adenyl cycl sulphate (APS)-kinase from Arabidopsis thaliana
JOURNAL	Biochim. Biophys. Acta 1218 (3), 447-452 (1994)
MEDLINE	94325358
PUBMED	8049272
REMARK	(sites)
REFERENCE	2 (bases 1 to 1185)
AUTHORS	Arz, H.E., Gisselmann, G., Schiffmann, S. and Schwenn, J.D.
TITLE	A chloroplast APS-Kinase cDNA from Arabidopsis thaliana
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 1185)
AUTHORS	Schwenn, J.D.
TITLE	Direct Submission
JOURNAL	Submitted (03-NOV-1993) Schwenn J. D., Ruhr University Bochum, Biology, Universitaetstr. 150, 44780 Bochum, Germany
FEATURES	Location/Qualifiers

FEATURES

```

1. 1185
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/db_xref="taxon:14CDNA"
/clone="PM239x14CDNA"
/tissue_type="leaf"
/clone_lib="cDNA in lambda ZapII"
/dev_stage="vegetative"
42. .872
/EC_number="2.7.1.25"
/functions="phosphorylates 3'-OH group of adenylylsulfate"
/standard_name="ATP:adenylylsulfate-3'-phosphotransferase"
/citations[2]
/codon_start=1
/product="APS kinase"
/protein_id="CAA53436.1"
/db_xref="GI:414737"
/db_xref="GOA:Q43295"
/db_xref="UniProt/Swiss-Prot:Q43295"
/tranlation="MTAAGAKSLGLLSMASPKGIFDPSNRSVVVVVACVSMDS
QTLSHNKGSIPEVKSINGHTGQKQSLPVSSTNWKHECSVEKVDRORLDDQKGC
VIWVTGUSGGKSLCANQMLYQKGLCYILDGNNVRHGLANDLSFKAEADRAENIR
RVEYAKLFDAGIGICTASITSPYTRDACSLLPDEGDFVFMVPLSVSCVARDPK
GLVGLAPRAGIKGTGIDDDPEYPLNCEISILGREGGTSPIEMAEKVGVYLDNKGYLQA

```

mat peptide

```

/product="APS-kinase"
/EC number="2.7.1.25"
/note="putative mature peptide only - gene product
starting with ATG at position 162 is enzymatically active,
but maturation in the chloroplast has not yet been shown.
Homology to sequences from E. coli & ye"
/citation={1}
/citation={2}

```

transit peptide

```

stem_loop
polyA_site
stem_loop
984. .1012
985. .990
1092. .1120
Let. 261, 455, 1990"
/notes="putative transit peptide with a processing site
suitable for stromal protease; ref. Gavel & Heijne FEBS

```

ORIGIN

Alignment Scores:		
Pred. No.:	1.13e-30	1185
Score:	816.00	159
Percent Similarity:	77.8%	317
Best Local Similarity:	65.16%	44
Query Match:	44.69%	10
DB:	8	3
		Gaps:
		Indels:
		Mismatches:
		Conservative:
		Matches:
		Length:

US-10-829-432-4 (1-343) x ATCLAPSK (1-1185)

Qy	100	GlyAlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgArgAsn	119
Db	168	GGATCTCAAACCTCTGAGTCATAACAAATAATGGATCTATTCTCGAGGTAAATCCATTAAC	227
Qy	120	GlyAlaAlaProGlyGluAlaProHisHisSerProValIlysGluIysProValMetSerAsn	139
Db	228	GGT-----CACACGGGACAAAGCAAGCAACT---TTGTCTTACG	263
Qy	140	IleGlyIysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln	159
Db	264	GTCCGAACCTCGACAAATATAAAGTGGCATGATGTTCTGTCAGAAAGTTGATACAG	323
Qy	160	LysLeuLeuGlyGlnIlysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys	179
Db	324	AGATTGCTTGATCAGAAAGGATGTGTGTTGGGTACCGCTCTTAGTGTTCAGGGAAG	383
Qy	180	SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal	199
Db	384	AGTACTTTTGCTTGTGCTTTGTAATCAGATGTTGTATCAAAGGGGAAGCTTTGTTATATT	443
Qy	200	LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheTysAlaGluAsp	219
Db	444	CTTGATGTCATAATGTTAGGCATGGCTTAAACCGTGATCTTAGCTTTAAAGCTGAGGAT	503
Qy	220	ArgAlaGluAsnIleArgArgValGlyGluValAlaIlysLeuPheAlaAspAlaGlyVal	239
Db	504	CGTGACAGAAATATTCGTAGAGTTGGAGAGGTTGCTAAAGCTTTTCCGGATGCTGAATA	563
Qy	240	IleCysIleAlaSerLeuIleSerProTyrArgAspAspAspAlaCysArgAlaLeu	259
Db	564	ATCTGCATTGCGAGTTTGATATCTCTTATAGAACAGATAGGACGCTTGTCTGAAGTTTG	623
Qy	260	LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuIysIleCysGluAla	279
Db	624	CTCCCCGAGGAGATTTTGTGTAGGTGTTTCATGGATGTACCGCTTAGTGTTCGCGAGCG	683
Qy	280	ArgAspProIysGlyLeuTyrLysLeuAlaArgThrGlyLysIleIysGlyPheThrGly	299
Db	684	AGGATCCAAAGGTCCTTACAGACTGCTCTGTCGAGAAAGATCAAGGTTTATCCGGG	743
Qy	300	IleAspAspProTyrGluProProIleAsnGlyGluIleValIleIysMetLysAspGlu	319
Db	744	ATCGATGACCTTACGAGCCACCATTGAACCTGGCAG-----ATTTCTTAGGACGCTGAA	797
Qy	320	GluCysProSerProIysAlaMetAlaLysGlnValLeuCystyrLeuGluGluAsnGly	339
Db	798	GGAGGAACCTTCTCTATCGAAATGGCGGAAAGGTCGTGGATACTTAGATATAAAGGGT	857
Qy	340	TyrLeuGlnAla	343
Db	858	TATCTTCAAGCA	869
RESULT 12			
AF178976		970 bp	linear
LOCUS		Zea mays adenosine-5'-phosphosulfate kinase (AK1)	PLN 13-DEC-1999
DEFINITION			cds
ACCESSION	AF178976		
VERSION	AF178976.1	GI:6563284	
KEYWORDS			
SOURCE			
ORGANISM			
			Zea mays
			Zea mays
			Embryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE			1 (bases 1 to 970)
AUTHORS			Bolchini, A., Petrucci, S. and Ottomello, S.
TITLE			Isolation and comparative expression analysis of a maize cDNA encoding adenosine 5'-phosphosulfate kinase
JOURNAL			Unpublished
REFERENCE			2 (bases 1 to 970)

AUTHORS Bolchi, A., Petrucco, S. and Ottonello, S.
TITLE Direct Submission
JOURNAL Submitted (19-AUG-1999) Istituto di Scienze Biochimiche,
 Universita' di Parma, Viale delle Scienze, Parma 43100, Italy
FEATURES Location/Qualifiers
 source 1..970
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="Dekalb XL72"
 /db_xref="taxon:4577"
 gene <1..970
 /gene="AK1"
 CDS <1..868
 /gene="AK1"
 /EC number="2.7.1.25"
 /note="APS kinase; ATP adenosine-5'-phosphosulfate
 3'-phosphotransferase"
 /codon_start=2
 /product="adenosine-5'-phosphosulfate kinase"
 /protein_id="AAFI17236.1"
 /db_xref="GI:563285"
 /translation="VPEPLLLAVARRRAPPEPRVKSRVASAPPPETPSGNQ
 NEDRRKLRGTPVLECTGDRSVEEQPEHAGVGKAWMSSTVPKSNSTVPHWDCPVKK
 TDQNVLRKQGVVWITGLSGSKSTLACGLRELHTRGKLAVYLDGDNLRHGLNDKL
 GFNAEDRAENRRVAKLFDAGLVCIASLISPHRRDRSCRALLSDDSIYEVFLNMS
 LEICEARDPKGLKYLARAGIKGTGTGIDDPYEAPLNCEIEIKVDVGVCPPPAEMAGGV
 VTYLEEKGFLEH"

ORIGIN

Alignment Scores:			
Pred. No.:	9.78e-31	Length:	970
Score:	815.50	Matches:	178
Percent Similarity:	67.97%	Conservative:	30
Best Local Similarity:	58.17%	Mismatches:	65
Query Match:	44.66%	Indels:	33
DB:	8	Gaps:	6

US-10-829-432-4 (1-343) x AF178976 (1-970)

QY	40	ProThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaProProValLeuPro	59
DB	33	CCCGCGCTCGGTGCGCGACCTG-----AACCGAGCCCGCGTCAATCTA	80
QY	60	GlyLeuThrProSerAspAlaProLeuProAlaLeuValIleHisGlyLeuThrProArg	79
DB	81	GGGTTCATCCGGCGCGCCCTCTCCCGAGACTCCAGCGGTGGACCAACGAGG	140
QY	80	SerSerHisSerSerAlaGlyLeuAlaSerAspSerGlyArgArgGluGlyGluArg	99
DB	141	GTG-----ACAGCGCACGAAAGCTCGAGGGACGC	170
QY	100	GlyAla-ArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgArgAs	119
DB	171	CCCTCGTCGAATGCAC-----TGGCGCCGTTCCTCGTCCGA-----	204
QY	119	nGlyAlaAlaProGlyGluAla-ProHisSerProValLysGluLys-----ProValM	137
DB	205	-----GGACACCGCGGAGCACGAGGAGTTGAGGGAAAGCTTGGAATGT	251
QY	137	etSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyClnSerA	157
DB	252	CATCCACCGTGGCGAAGTCATCAAAATATCTTCGGCACGATTTGCCAGTTGGCAAGACTG	311
QY	157	spArgGlnLysLeuLeuGlyGlnLysGlyCysValValTTrpIleThrGlyLeuSerGlyS	177
DB	312	ATGCCAAGACGCTACTCAAGCAGAAAGCGCTGTGTGTCTGATTACAGGCCCTTAGTGGTT	371
QY	177	erGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeu	197
DB	372	CAGTAAAGTACCTTGGCATGTACATTAGCCGTGAGCTCCACACAGAGGAAAGCTTG	431
QY	197	hrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysA	217

US-10-829-432-4 (1-343) x BD248389 (1-920)

```
QY      40  ProThrProThrLeuAlaValIleLeuValAenProGlnAqAlaProProVal----- 57
Db      14  CCGTCGGAATCAATGAAGATTAGCTATACAGAGCATCTCGA---CCGTCGGTITTTCTGT 70
QY      58  ---LeuProGlyLeu-----ThrProSerAspAla-----Pro 67
Db      71  TCTATTCAGGTCCTCGCGCGGATTCACCAACCACTCCAAAGTCAGGGTTCCTCAAG 130
QY      68  LeuProAlaLeuValIleHisGlyLeuThrProAArgSerSerHisSerAlaGlyLeu 87
Db      131  CTGCTCGGTC-GTCTATTCCGGCGGACAGCCGCAAAATTAGTAGCAATTTCTACTCTCT 189
QY      88  AlaSerAspSerGlyArgGluGlyLeuGlyArgGlyAlaAArgThrHisCysHisArg 107
Db      190  TCATCAATCTCAGCCGT-----TAACTGTCTCTGCTCAAGTCTCCCTCACCGC 237
QY      108  GlyIleGlyArg-----TrpValArgArg-ArgArgArgAsnGlyValaAlaPr 123
Db      238  TGAATTTCCCGCCCTTCAGAACTATCTAGAAAGGGGAGAAATAACGGA----- 289
QY      123  oGlyGluAlaProHisSerProValLysGluLysProValMetSerAsnIleGlyLysSe 143
Db      290  -----AAAGAG----- 300
QY      143  rThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuGlu 163
Db      301  AGAGACATCTGTGTGCGACGAGATTCGATATGCGAGTCGACACACAACTCTTCA 360
QY      163  yGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAl 183
Db      361  ACAAAAGGGTGTGTGCTGTGATCTACTGTCTCAGTGTCTCAGGAAAGACACTGTTC 420
QY      183  aCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrrValLeuAspGlyAs 203
Db      421  TTGTGCACCTAAGTAAAGCATTTGTTTGAAGAGGCAAACTTACTTACACACTCGACGGCA 480
QY      203  pAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAs 223
Db      481  CAATGTCGTCACGGCTTTAACCGTGACCTCACTTCAAGCAGAGCACCAGCCGCAAAA 540
QY      223  nIleArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAl 243
Db      541  CATTAGAAGAAATTTGTGAGGTGGCTAAGTTGTTTGTCTGACGTCGAGTCAATTTGTATAGC 600
QY      243  aSerLeulleSerProTyrrArgArgAspArgAspAlaCysArgAlaLeuLeuProHisSe 263
Db      601  AAGTTTGATTTCTCCGTACCGGAGACACAGACGCGTCCCGTCTCTTTACCTGACGG 660
QY      263  rAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProly 283
Db      661  CGATTTCTGCGAGGCTTTCATGCACTTCTCTTCTGTCGCGAGTCGAGATCCAAA 720
QY      283  sGlyLeuTyrrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAspPr 303
Db      721  GGGGTTGTACAAGCTTTCACGTCGAGGCAAAATCAAAAGCTTCACTGGAATCGACGACC 780
QY      303  oTyrrGluProProIleAsnGlyGluIleValIleLys---MetLysAspGluGluCysPr 322
Db      781  TTACGAGGCACCATGTAATTTGCGAGGTAGTGTCTGAAACACACAGAGACGACGATCGTG 840
QY      322  oSerProLysAlaMetAlaLysGlnValLeuLysCysTyrrLeuGluAsnGlyTyrrLeuG 342
Db      841  TTCGCCACGTCAGATGGCTGAGAAATCATCTCTTACCTGCAAAACAAAGGTATCTTGA 900
QY      342  n 342
Db      901  G 901
```

RESULT 14

AF462823

LOCUS

1243 bp mRNA linear PLN 08-JAN-2002

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Arabidopsis thaliana AT4g39940/T5J17_110 mRNA, complete cds.
AF462823
AF462823.1 GI:18087562
FLI CDNA.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
AUTHORS

1 (bases 1 to 1243)
Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L.,
Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
and Ecker,J.R.

TITLE
JOURNAL

Arabidopsis cDNA clones
Unpublished
2 (bases 1 to 1243)

REFERENCE
AUTHORS

Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L.,
Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,
Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M.,
Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M.,
Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K.,
Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A.
and Ecker,J.R.

TITLE
JOURNAL

Direct Submission
Submitted (27-DEC-2001) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN
Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H.,
Kim,C.J., Meyers,M.C., Shinn,P., Banh,J., Bowser,L., Dale,J.M.,
Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G., Lam,B.,
Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M., Onodera,C.S.,
Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura, Y., Yu,G., Yu,S., Davis,R.W.,
Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES
source

1..1243
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clones="RAFL09-09-101(R21470)"
/ecotype="Columbia"
1..69
70..951
/note="adenosine-5-phosphosulfate-kinase"

5'UTR
CDS

/codon_start=1
/product="AT4g39940/T5J17_110"
/protein_id="RAL58913.1"
/db_xref="GI:18087563"
/translation="MEGLAINAKSRPSVFCSTPGLGGSHRKPDDGFLKLPASSIPAD
SRKLVA NSTSPFPI SAVNVAQSLTADFPALSETILKEGNNKGKAENIVHESLI
CRCDROLQKQGVVITGLSGSKSTVACALKALFERKGLTYTIDGNDVHGLNLR
DLTFKAERTNIRRI GEVAKLFADVGVICIASLISPYRRDRDACRSLLPDGFVEVFR
MDVPLHWCESDPKGLYKLARAGKIGKFTGIDDDPYEAPVNCVVLKHTGDDDESCPRQ
MAENIIISLQNKGYLEG"

```

3' UTR      952..1243
ORIGIN
Alignment Scores:
Pred. No.:      7,27e-29      Length:      1243
Score:          778.50      Matches:      175
Percent Similarity: 66.04%      Conservative: 37
Best Local Similarity: 54.52%      Mismatches: 67
Query Match:      42.63%      Indels:      43
DB:              8          Gaps:      9

US-10-829-432-4 (1-343) x AF462823 (1-1243)
QY      40  ProThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaProVal----- 57
Db      58  CCGTCGGATTCAATGAAGATTAGCTATCAGAGCATCTCGA---CCGTCGGTTTCTGT 114
QY      58  ----LeuProGlyLeu-----ThrProSerAspAla-----Pro 67
Db      115 TCTATTCCAGGTCCTCGCGCGGATTCCACCGAAACCTCCAAAGTGACGGTTTCCTCAAG 174
QY      68  LeuProAlaLeuValIleHisGlyLeuThrProArgSerSerHisSerAlaGlyLeu 87
Db      175  GTGCTCGCGTC-GTCTATTCCGCGGACAGCCGAAATATTAGTAGCGAATTTCTACTTCCTT 233
QY      88  AlaSerAspSerGlyArgArgGlyGlyGlyArgGlyAlaArgThrHisCysHisArg 107
Db      234  TCATCCAAATCTACGCGT-----TACGTCCTCTGCTCAAGCTTCCTCACCOC 281
QY      108 GlyIleGlyArg-----TrpValArgArg-ArgArgArgAsnGlyAlaAlaPr 123
Db      282  TGATTTTCCCGCCCTTTACAGAACTATATCTGAAGAGAGGAGGAATAACCGA----- 333
QY      123  cGlyGluAlaProHisSerProValGlyGlyLysProValMetSerAsnIleGlyLysSe 143
Db      334  -----AAAGAG-----AAAGC 344
QY      143  rThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuGlu 163
Db      345  AGAGACATCGTGTGCGACAGAGTTCGATATCGAGATCGACACACACACTTCTTCA 404
QY      163  yGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAl 183
Db      405  ACAAAAGGGTGTGTGCTTGGATCCTCTCAGTGTGTTTCTAGGCGGAGAAAGCACTGTTGC 464
QY      183  aCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAspGlyAs 203
Db      465  TTGTGCACCTAAGTAAGCACTTGTGTAAGAGGCGCAAACTTACTTACACACTCGACGCGCA 524
QY      203  pAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAs 223
Db      525  CAATGTCGCTACGGCCTTAACCGTGACCTCACTTCAAGCAGAGACCCGACCCAGAAA 584
QY      223  nIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAl 243
Db      585  CATTAGAACAATTTGTGAGTGGCTAAAGTTGTTTGTCTAGCGTCGAGTCAATTGTATAGC 644
QY      243  aSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeuLeuProHisSe 263
Db      645  AAGTTTGATTCTCCGTACCGAGAGACAGACGCGTCCCGCTCTGTTTACCTGACGG 704
QY      263  rAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProly 283
Db      705  CGATTTTCGTGAGAGTCTTCATGAGAGTCTCTCTTCATGTGCGGAGTCAGAGATCCAAA 764
QY      283  sGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAspPr 303
Db      765  GGGTTGTACAAGCTTTCAGCTGACGAGGCAAAATCAAAAGCTTCACTGGAATCGACGCC 824
QY      303  oTyrGluProProIleAsnGlyGluIleValIleLys---MetLysAspGluGluCysPr 322
Db      825  TTACGAGGCACCAAGTGAATTCGAGGTAGTGTGTAACACACACAGGAGACGACGAGTCGTG 884

```

322 oSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGlyTyrLeuGlu 342

885 TTCGCCAGCTCAGATGGCTGAGAACATCATCTCTTACCTGCAAAACAAAGGTTATCTTGA 944

342 n 342

945 G 945

RESULT 15

AF043351

LOCUS

DEFINITION

Arabidopsis thaliana adenosine-5'-phosphosulfate-kinase (akn2)

VERSION

AF043351.1

GI:2829132

KEYWORDS

Arabidopsis thaliana (chale cress)

ORGANISM

Arabidopsis thaliana

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE

1 (bases 1 to 1311)

Arz,H.B., Gisselmann,G., Schiffmann,S. and Schwenn,J.D.

A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana

Biochim. Biophys. Acta 1218 (3), 447-452 (1994)

94325358

8049272

2 (bases 1 to 1311)

Schiffmann,S. and Schwenn,J.D.

Isolation of cDNA clones encoding

adenosine-5'-phosphosulfate-kinase (BC2.7.1.25) from Catharanthus

roseus (Accession No. AF044285) and an isoform (akn2) from

Arabidopsis (Accession No. AF043351) (PGR98-116)

Plant Physiol. 117 (3), 1125 (1998)

3 (bases 1 to 1311)

Schiffmann,S. and Schwenn,J.D.

Direct Submission

Submitted (16-JAN-1998) Biochemie der Pflanzen, Ruhr-Universitaet

Bochum, Universitaetsstrasse 150, Bochum 44780, Germany

Location/Qualifiers

1..1311

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/cultivar="Columbia"

/db_xref="taxon:3702"

/clone="PGAD10"

/tissue_type="green vegetative"

/dev_stage="3 week old"

1..1311

/gene="akn2"

61..942

/gene="akn2"

/EC_number="2.7.1.25"

/function="phosphorylates 3'-OH group of

adenosine-5'-phosphosulfate"

/note="APS-kinase isoform; similar to the product encoded

by Genbank Accession Number X75782"

/codon_start=1

/product="adenosine-5'-phosphosulfate-kinase"

/protein_id="AAC39520.1"

/db_xref="GI:2829133"

/translation="MEGLAIRASRPSVFCSTPGLGDSHRKPPSPDGLKLPASSIPAD

SRKLVANSTSFHPISAVNVSAQASLTADFPALSETILKEGRNGKEKAENIVWHESSI

CRDRQQLQOQGVVMTIGUSGSGKSTVACALSKALFPERKLTITLQDGNVRHGLNR

DLTPFAEHTENIRRIEVAKLFPADVGVICSLISPYRRDRDRACRSLLPDGFVEVF

MDVPLHVCESRDPKGLYKLRAGAKIKGTGTDDPFEAPVNCVFLVLTGDDSCSPRQ

MAENIISYLQNKVLEG"

238..240

/misc_feature

238..240

/gene="akn2"

/note="putative processing site 3' of this position;

unclassified site"

ORIGIN

Alignment Scores:

Pred. No.: 7,67e-29 Length: 1311
Score: 778.50 Matches: 175
Percent Similarity: 66.04% Conservative: 37
Best Local Similarity: 54.52% Mismatches: 67
Query Match: 42.63% Indels: 43
DB: 8 Gaps: 9

US-10-829-432-4 (1-343) x AF043351 (1-1311)

```
QY 40 ProThrProLeuAlaValLeuValAsnProGlnArgAlaProProVal----- 57
   |||:::  ::::  :::  |||
Db 49 CCGTCGGATTCAATGAGGATTAGCTATCAGAGCATCTCGA---CCGTCGGTTTCTGT 105
   |||:::  |||  |||  |||
QY 58 ---LeuProGlyLeu-----ThrProSerAspAla-----Pro 67
   |||:::  |||  |||  |||
Db 106 TCTATTCCAGGTCTCGCGCGGATTCCACCGAAACCTCCAAGTGACGGTTTCTCAAG 165
   |||:::  |||  |||  |||
QY 68 LeuProAlaLeuValIleHisGlyLeuThrProArgSerSerHisSerSerAlaGlyLeu 87
   |||:::  |||  |||  |||
Db 166 CTGCTGGGTC-GTCTATTCCGGCGGACAGCCGAAATATTAGTAGCGAATTCTACTTCCTT 224
   |||:::  |||  |||  |||
QY 88 AlaSerAspSerGlyArgArgGluGlyGluGlyAlaArgGlyAlaArgThrHisCysHisArg 107
   |||:::  |||  |||  |||
Db 225 TCATCCAATCTCAGCCGT-----TAACGTCTCTGCTCAAGCTTCCCTCACCGC 272
   |||:::  |||  |||  |||
QY 108 GlyIleGlyArg-----TrpValArgArg-ArgArgArgAsnGlyAlaAlaPr 123
   |||:::  |||  |||  |||
Db 273 TGATTTTCCCGCCCTTTCAGAACTACTGAAAGAGGAAGAAATAACGGA----- 324
   |||:::  |||  |||  |||
QY 123 oGlyGluAlaProHisSerProValIysGluIysProValMetSerAsnIleGlyLysSe 143
   |||:::  |||  |||  |||
Db 325 -----AAGAG-----AAGC 335
   |||:::  |||  |||  |||
QY 143 rThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuG1 163
   |||:::  |||  |||  |||
Db 336 ACAGAACATCGTGTGCACGAGATTGCGATATGCAGATGCGACAGACACAACTTCTTCA 395
   |||:::  |||  |||  |||
QY 163 yGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAl 183
   |||:::  |||  |||  |||
Db 396 ACAAAAGGGTTGTGCTTTGGATCACTGCTCTCAGTGGTTTCAAGGAAAGCAGCTGTGC 455
   |||:::  |||  |||  |||
QY 183 aCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrrValLeuAspGlyAs 203
   |||:::  |||  |||  |||
Db 456 TTGTGCACCTAAGTAAAGCATTTGTTGAAAGAGGCAAACTTACTTACACACTCGACGGCA 515
   |||:::  |||  |||  |||
QY 203 PheLeuArgHisGlyLeuAsnArgAspLeuSerPheIysAlaGluAspArgAlaGluAs 223
   |||:::  |||  |||  |||
Db 516 CAATGTCCGTACGGGCTTAAACCGTACCTCTTCAAAGCAGACGCCGCCACCGAANA 575
   |||:::  |||  |||  |||
QY 223 nIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAl 243
   |||:::  |||  |||  |||
Db 576 CATTAGAGAGANTGTGAGTGGCTAAGTTGTTTCTGCTGCGAGTCATTGTATAGC 635
   |||:::  |||  |||  |||
QY 243 aSerLeuIleSerProTyrrArgAspAlaCysArgAlaLeuLeuProHisSe 263
   |||:::  |||  |||  |||
Db 636 AAGTTTGTATTCTCCGTACCGAGAGACAGACGCGTCCGCTCTTGTACCTGACGG 695
   |||:::  |||  |||  |||
QY 263 rAsnPheIleGluValPheIleAspLeuProLeuIysIleCysGluAlaArgAspProly 283
   |||:::  |||  |||  |||
Db 696 CGATTTCGTGAGGGTCTTCATGGAGCTTCTTATGTGTGCGAGTCGAGAGATCCAAA 755
   |||:::  |||  |||  |||
QY 283 sGlyLeuTyrrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAspPr 303
   |||:::  |||  |||  |||
Db 756 GGGGTTGTACAAGCTTGACGTGCAGGCAAAATCAAAGCTTCACTGGNATCGACGCC 815
   |||:::  |||  |||  |||
QY 303 oTyrrGluProProIleAsnGlyGluIleValIleLys---MetLysAspGluGluCysPr 322
   |||:::  |||  |||  |||
Db 816 TTACGAGGCCACCAAGTGAATTGCGAGGTAGTGTGAAACACACAGGAGACGACGAGTGGT 875
   |||:::  |||  |||  |||
QY 322 oSerProLysAlaMetAlaLysGlnValLeuCysTyrrLeuGluGluAsnGlyTyrrLeuG1 342
   |||:::  |||  |||  |||
Db 876 TTCGCCACGTGAGTGGCTGAGAACATCTCTTACTCTGCAAAACAAGGTTATCTTGA 935
   |||:::  |||  |||  |||
```

QY 342 n 342
Db 936 G 936

Search completed: September 9, 2005, 14:24:24
Job time : 4078 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 9, 2005, 13:00:56 ; Search time 531 Seconds
(without alignments)
3823.863 Million cell updates/sec

Title: US-10-829-432-4

Perfect score: 1826

Sequence: 1 RPFHINQTEPLVTHQQPP.....PKMAKQVLCYLENGYLQA 343

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10829432/runat_08092005_132222_25068/app.query.fasta_1.519
-DB=N Geneseq_16Dec04 -OFT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10829432 @CN_1_1_708 @runat_08092005_132222_25068 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_16Dec04:.*
1: Geneseqn1980s:.*
2: Geneseqn1990s:.*
3: Geneseqn2000s:.*
4: Geneseqn2001as:.*
5: Geneseqn2001bs:.*
6: Geneseqn2002as:.*
7: Geneseqn2002bs:.*
8: Geneseqn2003as:.*
9: Geneseqn2003bs:.*
10: Geneseqn2003cs:.*
11: Geneseqn2003ds:.*
12: Geneseqn2004as:.*
13: Geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1826	100.0	1217	AAZ50160	Corn Aden
2	887	48.6	890	AAZ50159	Corn Aden
3	876	48.0	928	AAZ50163	Wheat Ade
4	852	46.7	936	AAZ50162	Soybean A
5	834	45.7	627	AAZ44691	Arabidops

6	817	44.7	1175	3	AAC36588	Aac36588 Arabidops
7	816	44.7	831	6	ABZ12905	Abz12905 Arabidops
8	816	44.7	831	6	ADG87955	Adg87955 A. thalia
9	816	44.7	831	12	ADN74374	Adn74374 Thale cre
10	816	44.7	1068	3	AAC35266	Aac35266 Arabidops
11	816	44.7	1183	3	AAC45659	Aac45659 Arabidops
12	809.5	44.3	714	3	AAC51190	Aac51190 Arabidops
13	788	43.2	687	12	ADJ39869	Adj39869 Plant cDN
14	778.5	42.6	917	3	AAC298337	Aac298337 A. thalia
15	778.5	42.6	948	3	AAC42294	Aac42294 Arabidops
16	757.5	41.5	1130	3	AAC49070	Aac49070 Arabidops
17	752.5	41.2	1133	3	AAC36627	Aac36627 Arabidops
18	640	35.0	378	12	ADJ44283	Adj44283 Plant cDN
19	617.5	33.8	483	3	AAC51315	Aac51315 Arabidops
20	571	31.3	603	8	AAC27590	Aac27590 Prokaryot
21	561	30.7	595	12	ADJ43843	Adj43843 Plant cDN
22	555	30.4	600	13	ADT42933	Adt42933 Bacterial
23	534.5	29.3	609	13	ADT42467	Adt42467 Bacterial
24	521.5	28.6	597	8	ACA46934	Aca46934 Prokaryot
25	521.5	28.6	615	6	ABN92217	Abn92217 Staphyloc
26	521.5	28.6	615	13	ADS02000	Ads02000 Staphyloc
27	519.5	28.5	588	13	ADT41933	Adt41933 Bacterial
28	517	28.3	681	10	ADF02710	Adf02710 Bacterial
29	516	28.3	1863	13	ADS58748	Ads58748 Bacterial
30	514.5	28.2	591	8	ACA22457	Aca22457 Prokaryot
31	514.5	28.2	612	8	ACA44895	Aca44895 Prokaryot
32	514	28.1	28401	10	ADC97728	Adc97728 Rhizobito
33	510.5	28.0	1842	13	ADS63529	Ads63529 Bacterial
34	510.5	28.0	1869	13	ADS84272	Ads84272 Bacterial
35	510.5	28.0	1890	13	ADS63901	Ads63901 Bacterial
36	510	27.9	642	8	ACA53672	Aca53672 Prokaryot
37	505.5	27.7	597	8	ACA47863	Aca47863 Prokaryot
38	503.5	27.6	6715	13	ADR84259	Adr84259 Aspergill
39	503	27.5	603	8	ACA35030	Aca35030 Prokaryot
40	503	27.5	651	11	ACH97008	Ach97008 Klebsiell
41	502.5	27.5	2646	6	AAS94855	Aas94855 Human DNA
42	501	27.4	633	10	ACF69692	Acf69692 Photorhab
43	501	27.4	110000	10	ACF67367	26 Continuation (27 o
44	501	27.4	110000	10	ACF65386	2 Continuation (3 o
45	498.5	27.3	1983	13	ADT42849	Adt42849 Bacterial

ALIGNMENTS

RESULT 1
AAZ50160
ID AAZ50160 standard; cDNA; 1217 BP.
XX
AC AAZ50160;
XX
DT 04-MAY-2000 (first entry)
XX
DE Corn Adenyllysulphate kinase-2 cDNA clone.
XX
KW Adenyllysulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
3'-phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
XX
KW corn; clone p0016.ctscj40rb; transgenic plant; screen; antibody; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
CDS 2..1033
FT /*tag= a
FT /product= "Corn APS kinase-2"
FT /note= "Derived from clone p0016.ctscj40rb"
XX
PN WO200004165-A1.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 98WO-US015809.
XX
PR 14-JUL-1998; 98US-0092833P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX PA Falco SC, Allen SM, Anderson SL;
 XX PI WPI; 2000-182430/16.
 XX DR P-PSDB; AAY44789.
 XX PT New nucleic acid molecule and chimeric gene encoding an adenosine-5'
 PT phosphosulfate kinase, useful for altering expression of sulfate
 PT assimilation protein in plants.
 XX Claim 3; Page 30; 42pp; English.
 XX The present sequence is a cDNA encoding corn adenylylsulphate kinase (APS
 CC kinase), also known as adenosine-5' phosphosulphate kinase. This is
 CC obtained from clone p0016.ctscj40rb, derived from corn pooled tassel
 CC shoots, p0016 cDNA library. APS kinase is a sulphate assimilation
 CC protein, that catalyses the conversion of adenosine-5' phosphosulphate
 CC (APS) to 3'-phospho-adenosine-5' phosphosulphate (PAPS). This sequence is
 CC used as probe and primers to identify, obtain and synthesise sulphate
 CC assimilation proteins from other plants. It is also used to produce
 CC transgenic plants, that are useful for altering the expression levels of
 CC a sulphate assimilation protein. The APS kinase peptides are useful for
 CC producing antibodies, that are used to screen and isolate cDNA clones
 XX SQ Sequence 1217 BP; 344 A; 292 C; 303 G; 278 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.61e-103 Length: 1217
 Score: 1826.00 Matches: 343
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-10-829-432-4 (1-343) x AAZ50160 (1-1217)

QY 1 ArgProPheHisPheIleAsnGlnThrGluProLeuValThrHisThrGlnGlnProPro 20
 DB 2 CGTCGGTTTCATTTTATCATCAACAGAACTCTGGTCACACACAGGACACACCG 61
 QY 21 SerProAlaProGlyProAlaSerGlnGlnGlnArgGlnGlnGlnGlnGlnGlnGln 40
 DB 62 AGCCCGCGCCCGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 121
 QY 41 ThrProThrLeuAlaValIleValLeuValLeuValLeuValLeuValLeuValLeu 60
 DB 122 AGCCCGAGCGTCGCGTCATCTCGTAAATCCACAGCGCGCGCTCCCGTCTCCAGGC 181
 QY 61 LeuThrProSerAspAlaProLeuProAlaLeuValIleHisGlyLeuThrProArgSer 80
 DB 182 CTCACCCCTAGGATGCGGCATCTCCGGGCTCGTGATCTCATGGGCTACTCCCGTTCC 241
 QY 81 SerHisSerSerAlaGlyLeuAlaSerAspSerGlyArgArgGluGluGluGlyArgGly 100
 DB 242 TCACACTCTTCGCGGCTCTGCCAGTGATGATGCGCGCGCGCGAGGGGAGGCGCGGT 301
 QY 101 AlaArgThrHisCysHisArgGlyIleGlyArgGlyValArgArgArgArgArgArg 120
 DB 302 GCGCGTACGACATGCGCGCATTTGGCGGCTGGTGGCGCGCGCGCGCGGAGTGA 361
 QY 121 AlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsnIle 140
 DB 362 GAGCGCGCGGAGGCGCGCGCACAGCCAGTGAAGGAGAGCCCTGTATGTGACAT 421
 QY 141 GlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLys 160
 DB 422 GGGAAATCGACTAATATTTATGGACAAATTCCTGATGGACAATCTGATAGACAGAAA 481
 QY 161 LeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlyGlySer 180
 DB 482 TTGCTGGGACAAAGAGCGTGTCTGTATGATGATGATGATGATGATGATGATGATGAT 541

QY 181 ThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeu 200
 DB 542 ACTCTTGATGTCGACTGAGTCTGAGTTGATTCGAGAGGCCACCTCAGTATGACTT 601
 QY 201 AspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLeuAlaGluAspArg 220
 DB 602 GATGGTGACAACCTCAGACATGGCTAAATAGAGATTTAAGCTTTAAGGCAGAGACCGT 661
 QY 221 AlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIle 240
 DB 662 GCAGAAATATACGAAGAGTTGGTCAAGTGGCAAGCTTTTGTGATGCTGGGTGTCATA 721
 QY 241 CysIleAlaSerLeuLeuSerProTyrArgArgAspArgAspAlaCysArgAlaLeuLeu 260
 DB 722 TGCAATTGCTAGCTTGATATCTCCATACAGGAGAGATCGTGATGATGATGCTGCTACTT 781
 QY 261 ProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArg 280
 DB 782 CCACATTCTAACTTTATTGAAAGTATTTATTGATTTGCCCTTAAAAATTTGTGAAGCTCGT 841
 QY 281 AspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIle 300
 DB 842 GATCTTAAAGGCTTATACAGCTTGCAGTACAGGAAAGATTAAGAGTTTCACTGGAATT 901
 QY 301 AspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGluGlu 320
 DB 902 GATGATCCATACGAACCAACCAATTAATGTCAGATAGTAATTAAGATCAAGATGAGGAA 961
 QY 321 CysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGlyTyr 340
 DB 962 TGCCCTTCACCAAGCAATGCCAAGCAAGTTCTATGCTACCTTGAAGAAAAACGATAT 1021
 QY 341 LeuGlnAla 343
 DB 1022 TTGCAAGCT 1030

RESULT 2
 AAZ50159
 ID AAZ50159 standard; cDNA; 890 BP.
 XX AAZ50159;
 AC AAZ50159;
 DT 04-MAY-2000 (first entry)
 XX Corn adenylylsulphate kinase-1 cDNA clone.
 DE Adenylylsulphate kinase; Adenosine-5'-phosphosulphate kinase; APS kinase;
 KW 3'-Phospho-adenosine-5'-phosphosulphate; PAPS; sulphate assimilation;
 KW corn; clone cen3n.pk0088.b10; transgenic plant; screen; antibody; ss.
 OS Zea mays.
 XX
 FH Key Location/Qualifiers
 CDS 3..677
 FT /*tag= a
 FT /product= "Corn APS kinase-1"
 FT /note= "Derived from clone cen3n.pk0088.b10"
 XX WO200004165-A1.
 PN 27-JAN-2000.
 PD 13-JUL-1999; 99WO-US015809.
 PF 14-JUL-1998; 98US-0092833P.
 PR (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Falco SC, Allen SM, Anderson SL;
 XX WPI; 2000-182430/16.
 DR P-PSDB; AAY44788.


```
Db 460 CTGAGTCAAAGCTTGCACTCCAAAGGAAACTGCTTTACATCCTTGATGTCACATATT 519
Qy 206 AtgHieGlyLeuAenArgAspLeuSerPheLysAlaGluAspArgAlaGluAenIleArg 225
Db 520 CGGCATGGTCTTAACAGAGATCTTAGTTTGTAGAGCAGAGATGCTTCTGAACAACATTAGA 579
Qy 226 ArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeu 245
Db 580 AGGATTGGTGAGGTGGCAAACTCTTTGCAGATGCTGGTGTATTATTTGCATCATGATTTA 639
Qy 246 IleSerProTyrArgArgAspArgAspAlaCysArgAlaLeuLeuProHisSerAenPhe 265
Db 640 ATATCACTACCAATACCAAAAGATAGATGATGCAGACGACTACTTTCAAAGGAGATTTT 699
Qy 266 IleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProLysGlyLeu 285
Db 700 ATTGAGTTTTCATAGATGTTCCACTACATGTTGTGAAGCTAGGACCCAAAGGACATC 759
Qy 286 TyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAspProTyrGlu 305
Db 760 TACAAGCTTGCTCGAGCTGGAAGATCAAAAGGTTTCACTGGTATAGATCATCATATGAA 819
Qy 306 ProProIleAsnGlyGluIleValIleLysMetLysAspGluGluCysProSerProLys 325
Db 820 CCACCGTGTAGTTGTGAGATAGTATTACAAACAGAAAGGAAGTCACTGTAAAGTCTCCCACT 879
Qy 326 AlaMetAlaLysGlnValLeuCysTyrLeuGluGluAenGlyTyrLeuGlnAla 343
Db 880 GATATGGCTGAAGAGTATATCTTACTTGGAGGAGACGGATACCTGCGGGCT 933

RESULT 5
AAC44691
ID AAC44691 standard; DNA; 627 BP.
AC AAC44691;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 43768.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 08-APR-1999; 99US-0128234P.
XX 16-APR-1999; 99US-0128714P.
XX 19-APR-1999; 99US-0129845P.
XX 21-APR-1999; 99US-0130077P.
XX 23-APR-1999; 99US-0130449P.
XX 28-APR-1999; 99US-0130510P.
XX 30-APR-1999; 99US-0130891P.
XX 04-MAY-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 04-MAY-1999; 99US-0132407P.
XX 05-MAY-1999; 99US-0132484P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
XX 24-JUN-1999; 99US-0140695P.
XX 28-JUN-1999; 99US-0140823P.
XX 29-JUN-1999; 99US-0140991P.
XX 30-JUN-1999; 99US-0141287P.
XX 01-JUL-1999; 99US-0141842P.
XX 01-JUL-1999; 99US-0142154P.
XX 02-JUL-1999; 99US-0142055P.
XX 06-JUL-1999; 99US-0142390P.
XX 08-JUL-1999; 99US-0142803P.
XX 09-JUL-1999; 99US-0142920P.
XX 12-JUL-1999; 99US-0142977P.
XX 13-JUL-1999; 99US-0143542P.
XX 14-JUL-1999; 99US-0143624P.
XX 15-JUL-1999; 99US-0144005P.
XX 16-JUL-1999; 99US-0144085P.
XX 16-JUL-1999; 99US-0144086P.
XX 19-JUL-1999; 99US-0144325P.
XX 19-JUL-1999; 99US-0144331P.
XX 19-JUL-1999; 99US-0144332P.
XX 19-JUL-1999; 99US-0144333P.
XX 19-JUL-1999; 99US-0144334P.
XX 19-JUL-1999; 99US-0144335P.
XX 20-JUL-1999; 99US-0144352P.
XX 20-JUL-1999; 99US-0144684P.
XX 20-JUL-1999; 99US-0144884P.
XX 21-JUL-1999; 99US-0144814P.
XX 21-JUL-1999; 99US-0145086P.
XX 21-JUL-1999; 99US-0145088P.
XX 22-JUL-1999; 99US-0145085P.
XX 22-JUL-1999; 99US-0145087P.
XX 22-JUL-1999; 99US-0145089P.
XX 22-JUL-1999; 99US-0145192P.
```

```
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148315P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.

PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 25-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 8 2e-43
Score: 834.00
Percent Similarity: 88.35%
Best Local Similarity: 74.76%
Query Match: 45.67%
DB: 3
Gaps: 0

US-10-829-432-4 (1-343) x AAC44691 (1-627)

Qy 137 MetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyInSer 156
Db 1 ATGTCGACAGTGGGAAATTCACGACATATTTGGCAAGAATCCCCCATTTGGGAAACT 60

Qy 157 AspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGly 176
Db 61 GAAAGGCAGAAAGTTGTTAAATCAGAAAGGTTGTGTGTGTGGATCACAGGGCTCAGTGGC 120

Qy 177 SerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeu 196
Db 121 TCAGGGAAAGCAGCTAGCTTGTCTTGTAGAGAGCTGAAACACCGGGGAAAGCTA 180

Qy 197 ThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLys 216
Db 181 TCATATATTCTTGATGGTGACAACTCTGTCATGTTTGAACAAAGATCTTTGGTTTCAAG 240

Qy 217 AlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAsp 236
Db 241 GCAGAGGATAGATGGGAAATATACGAGGTCGGAGAGTAGCCAAACTCTTTGCGGAT 300

Qy 237 AlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspAspAlaCys 256
Db 301 CTGTTTAACTGTATTGCCAGCTCATATCCCGTATAGAAAGACCGTGACGCTGC 360

Qy 257 ArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIle 276
Db 361 AGGGAATGATACAGAATTCATCTTTATTTCAGGTTTTCATGAACATGCTCTGCAATTG 420

Qy 277 CysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGly 296
Db 421 TGTGAAGCAAGGGACCCCTAAAGGCTATACAAAGCTTGACGTGCGAGGAAGATCAAGGT 480

Qy 297 PheThrGlyIleAspAspProTyrGluProPheIleAsnGlyGluIleValIleLysMet 316
Db 481 TTCACAGGAATAGATGATCCGTATGAGTCTCCCTTGAATCTGTAGATAGATTGAAAGAG 540

Qy 317 LysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGlu 336
Db 541 AAAGAGGGAGAGTGTCTTCCCTGTAGCTATGCTGAGGAAGTATCTCTTATCTAGAA 600

Qy 337 GluAsnGlyTyrLeuGln 342
Db 601 GACAAAGGTTTCTCTCAA 618

RESULT 6
AAC36588
ID AAC36588 standard; DNA; 1175 BP.
XX
AC AAC36588;
```

XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 14353.
DE
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
XX 09-MAR-1999; 99US-0123348P.
PR 21-JUN-1999; 99US-0139817P.
XX 23-MAR-1999; 99US-0125788P.
PR 22-JUN-1999; 99US-0139899P.
XX 25-MAR-1999; 99US-0126264P.
PR 23-JUN-1999; 99US-0140353P.
XX 29-MAR-1999; 99US-0126785P.
PR 23-JUN-1999; 99US-0140354P.
XX 01-APR-1999; 99US-0127462P.
PR 24-JUN-1999; 99US-0140695P.
XX 06-APR-1999; 99US-0128234P.
PR 28-JUN-1999; 99US-0140823P.
XX 08-APR-1999; 99US-0128714P.
PR 29-JUN-1999; 99US-0140991P.
XX 16-APR-1999; 99US-0129845P.
PR 09-JUL-1999; 99US-0142920P.
XX 19-APR-1999; 99US-0130077P.
PR 08-JUL-1999; 99US-0142803P.
XX 21-APR-1999; 99US-0130449P.
PR 12-JUL-1999; 99US-0142977P.
XX 23-APR-1999; 99US-0130510P.
PR 13-JUL-1999; 99US-0143542P.
XX 23-APR-1999; 99US-0130891P.
PR 14-JUL-1999; 99US-0143624P.
XX 28-APR-1999; 99US-0131449P.
PR 15-JUL-1999; 99US-0144005P.
XX 30-APR-1999; 99US-0132048P.
PR 16-JUL-1999; 99US-0144085P.
XX 04-MAY-1999; 99US-0132407P.
PR 19-JUL-1999; 99US-0144332P.
XX 05-MAY-1999; 99US-0132485P.
PR 19-JUL-1999; 99US-0144333P.
XX 06-MAY-1999; 99US-0132486P.
PR 19-JUL-1999; 99US-0144334P.
XX 07-MAY-1999; 99US-0132863P.
PR 19-JUL-1999; 99US-0144335P.
XX 11-MAY-1999; 99US-0134256P.
PR 20-JUL-1999; 99US-0144352P.
XX 14-MAY-1999; 99US-0134218P.
PR 20-JUL-1999; 99US-0144353P.
XX 14-MAY-1999; 99US-0134219P.
PR 20-JUL-1999; 99US-0144354P.
XX 14-MAY-1999; 99US-0134221P.
PR 20-JUL-1999; 99US-0144355P.
XX 18-MAY-1999; 99US-0134370P.
PR 21-JUL-1999; 99US-0144356P.
XX 19-MAY-1999; 99US-0134941P.
PR 21-JUL-1999; 99US-0144357P.
XX 20-MAY-1999; 99US-0135124P.
PR 21-JUL-1999; 99US-0144358P.
XX 21-MAY-1999; 99US-0135353P.
PR 22-JUL-1999; 99US-0144359P.
XX 24-MAY-1999; 99US-0135629P.
PR 22-JUL-1999; 99US-0144360P.
XX 25-MAY-1999; 99US-0136021P.
PR 22-JUL-1999; 99US-0144361P.
XX 27-MAY-1999; 99US-0136392P.
PR 23-JUL-1999; 99US-0144362P.
XX 28-MAY-1999; 99US-0136782P.
PR 23-JUL-1999; 99US-0144363P.
XX 01-JUN-1999; 99US-0137222P.
PR 23-JUL-1999; 99US-0144364P.
XX 03-JUN-1999; 99US-0137528P.
PR 26-JUL-1999; 99US-0145276P.
XX 04-JUN-1999; 99US-0137502P.
PR 27-JUL-1999; 99US-0145277P.
XX 07-JUN-1999; 99US-0137724P.
PR 27-JUL-1999; 99US-0145278P.
XX 08-JUN-1999; 99US-0138094P.
PR 27-JUL-1999; 99US-0145279P.
XX 10-JUN-1999; 99US-0138540P.
PR 28-JUL-1999; 99US-0145280P.
XX 10-JUN-1999; 99US-0138847P.
PR 28-JUL-1999; 99US-0145281P.
XX 14-JUN-1999; 99US-0139119P.
PR 28-JUL-1999; 99US-0145282P.
XX 16-JUN-1999; 99US-0139452P.
PR 29-AUG-1999; 99US-0145918P.
XX 17-JUN-1999; 99US-0139453P.
PR 09-AUG-1999; 99US-0147933P.
XX 18-JUN-1999; 99US-0139454P.
PR 10-AUG-1999; 99US-0147934P.
XX 18-JUN-1999; 99US-0139455P.
PR 10-AUG-1999; 99US-0147935P.
XX 18-JUN-1999; 99US-0139456P.
PR 11-AUG-1999; 99US-0148171P.
XX 18-JUN-1999; 99US-0139457P.
PR 12-AUG-1999; 99US-0148319P.
XX 20-AUG-1999; 99US-0149175P.
PR 13-AUG-1999; 99US-0148565P.
XX 20-AUG-1999; 99US-0149222P.
PR 13-AUG-1999; 99US-0148684P.
XX 20-AUG-1999; 99US-0149223P.
PR 16-AUG-1999; 99US-0149368P.
XX 20-AUG-1999; 99US-0149224P.
PR 17-AUG-1999; 99US-0149369P.
XX 23-AUG-1999; 99US-0149902P.
PR 20-AUG-1999; 99US-0149225P.
XX 18-JUN-1999; 99US-0139461P.
PR 23-AUG-1999; 99US-0149903P.
XX 18-JUN-1999; 99US-0139462P.
PR 25-AUG-1999; 99US-0150566P.
XX 18-JUN-1999; 99US-0139463P.
PR 26-AUG-1999; 99US-0150884P.

XX WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed and

XX producing plants with increased tolerance to these abiotic stresses.

XX Claim 144; SEQ ID NO 710; 577pp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid
 CC representative of expressed polynucleotides in the plant cell with an
 CC array or probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office
 XX
 SQ Sequence 831 BP; 224 A; 137 C; 223 G; 247 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.42e-41	Length:	831
Score:	816.00	Matches:	159
Percent Similarity:	77.87%	Conservative:	31
Best Local Similarity:	65.16%	Mismatches:	44
Query Match:	44.69%	Indels:	10
DB:	6	Gaps:	3

US-10-829-432-4 (1-343) x AB212905 (1-831)

QY	100	GlyAlaArgThrHisCysHisArgGlyIleGlyArgTnpValArgArgArgArgAsn	119
DB	127	GGATCTCAAACTGAGTCATACAAATAATGGATCTATTCCTGAGGTAAATCCATTAAC	186
QY	120	GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn	139
DB	187	GGT-----CACACGGGCAAAAGCAGGACCT---TTGCTACG	222
QY	140	IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln	159
DB	223	GTCGGAACCTCGACAAATATAAAGTGCATGAATGTTCTGTTGAGAAAGTTGATGACAG	282
QY	160	LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys	179
DB	283	AGATTGCTTGATCAGAAAGGATGTGTGATTTGGGTCCACCGGCTTAGTGTTTCAGGGAAG	342
QY	180	SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpVal	199
DB	343	AGTACTTGGCTTGCTTGATCATGATGTTGTATCAAAAGGGGAAGCTTTGTATATTT	402
QY	200	LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelLysAlaGluAsp	219
DB	403	CTTGATGGTGATAATGTTAGGCATGCTTAAACCGTGATCTTAGCTTTAAAGCTGAGGAT	462
QY	220	ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal	239
DB	463	CGTCGACAGAATATTCGTAGAGTTGAGAGGTTGCTTAAGCTTTTTCGGATGCTCGAATA	522
QY	240	IleCysIleAlaSerLeuIleSerProTrpArgArgAspAlaCysArgAlaLeu	259
DB	523	ATCTGATTCGCGAGTTTGATATCTCTTATAGAACAGATAGGACGCTTGTGAGATTTG	582
QY	260	LeuProHisSerAsnPhelIleGluValPheIleAspLeuProLeuLysIleCysGluAla	279
DB	583	CTCCCCGAGGAGATTTGTTGAGGTGTTTCATGATGTACCGCTTAGTGTTCGAGGGG	642
QY	280	ArgAspProLysGlyLeuTrpLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly	299
DB	643	AGGGATCCAAAGGGTCTTTACAAGCTTGTCTGTCGACGAGAAAGATCAAAAGGTTTACCGG	702

QY	300	IleAspAspProTrpGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu	319
DB	703	ATCGATGACCCCTTACGAGCCACCATTTGAACCTCGAG-----ATTCTCTAGGACGTGAA	756
QY	320	GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTrpLeuGluAsnGly	339
DB	757	GGAGAACTTCTCTATCGAAATGCGGAAGGTCGTCGATACCTAGATAACAAGGCT	816
QY	340	TyrLeuGlnAla	343
DB	817	TATCTTCAAGCA	828
XX	ADG87955	standard; cDNA; 831 BP.	
XX	ADG87955;		
XX	22-APR-2004	(first entry)	
XX	A. thaliana RPP7/RPP8-upregulated pathogen infection-related gene #397.		
XX	Pathogen infection-related gene; plant; Peronospora parasitica;		
XX	defence mechanism; RPP7; RPP8; pathogen resistance; transgenic plant;		
XX	oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.		
XX	Arabidopsis thaliana.		
XX	WO200222675-A2.		
XX	21-MAR-2002.		
XX	14-SEP-2001; 2001WO-US028506.		
XX	15-SEP-2000; 2000US-0232778P.		
XX	22-JUN-2001; 2001US-0300183P.		
XX	(SYGN) SYNGENTA PARTICIPATIONS AG.		
XX	(UYNC-) UNIV NORTH CAROLINA.		
XX	(GLAZ/) GLAZEBROOK J.		
XX	(WANG/) WANG X.		
XX	(DANG/) DANG J L.		
XX	(EULG/) EULGEM T.		
XX	(ZHUT/) ZHU T.		
XX	Glazebrook J, Wang X, Dangl JL, Eulgem T, Zhu T;		
XX	WPI; 2002-292409/33.		
XX	Novel isolated polynucleotide, useful for conveying pathogen resistance		
XX	to plants, and for identifying plants infected with a pathogen.		
XX	Claim 3; SEQ ID NO 397; 605pp; English.		
XX	The invention relates to 691 Arabidopsis thaliana genes (ADG87559--		
XX	ADG87557) whose expression is altered in response to pathogen infection,		
XX	and to homologues of these genes from other plants or fungi, especially		
XX	from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),		
XX	cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The		
XX	expression of genes of the invention was upregulated or downregulated in		
XX	Arabidopsis plants infected with the oomycete Peronospora parasitica,		
XX	indicating that they play a role in defence mechanisms. The genes of the		
XX	invention are regulated by RPP7 or RPP8 which act via unconventional		
XX	signalling cascades, or by the RPP4-dependent pathway. The invention also		
XX	relates to polypeptides encoded by the pathogen infection-related genes;		
XX	promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)		
XX	; expression cassettes, host cells and pathogen-resistant transgenic		
XX	plants and their progeny comprising a polynucleotide of the invention;		
XX	and a method of identifying a plant cell infected with a pathogen. The		
XX	polynucleotide sequences and methods of the invention are useful for		
XX	identifying plants infected with a pathogen, and for conferring		
XX	resistance to pathogens such as oomycetes, fungi, bacteria, viruses,		
XX	neomotodes and insects (e.g., aphids). The present sequence represents an		

CC Arabidopsis thaliana gene whose expression is altered in response to
CC Peronospora parasitica infection. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 831 BP; 224 A; 137 C; 223 G; 247 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,42e-41 Length: 831
Score: 816.00 Matches: 159
Percent Similarity: 77.8% Conservative: 31
Best Local Similarity: 65.16% Mismatches: 44
Query Match: 44.69% Indels: 10
DB: 6 Gaps: 3

US-10-829-432-4 (1-343) x ADG87955 (1-831)

QY 100 GlyAlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgArgAsn 119
DB 127 GGATCTCAAACTCTGAGTCATAACAAAATGGATCTATTCTCGAGGTAAATCCATTAC 186
QY 120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139
DB 187 GGT-----CACACGGGACAAAAGCAAGACCT---TTGTCTAGC 222
QY 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
DB 223 GTCGGAACCTCGACAATATATAAGTGGCATGAATGTCGTGAGAAAGTGTATAGACAG 282
QY 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
DB 283 AGATTGCTTGATCAGAAAGGATGTGTGATTGGGTCCACCGCTTTAGTGTTCAGGGAAG 342
QY 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpVal 199
DB 343 AGTACTTTGGCTTGCTTTGAATCAGATGTTGTATCAAAAGGGAAGCTTTGTTATATT 402
QY 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelysAlaGluAsp 219
DB 403 CTGTATGTTGATTAAGTGTAGGATGCTTAACCGTGATCTTAGCTTTAAAGCTGAGGAT 462
QY 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
DB 463 CGTGACAGAAATATTCGTAGAGTTGGAGAGTTGCTAAGCTTTTTCGGATCTGGAATA 522
QY 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspAspAlaCysArgAlaLeu 259
DB 523 ATCTGCATTGCGAGTTTGATATCTCCTTATAGAACAGATAGGACGCTTTGTCAAGTTTG 582
QY 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
DB 583 CTCCTCCAGGAGATTTTGAGGTGTCATGGATGACCCCTAGTGTTCGAGGCG 642
QY 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
DB 643 AGGGATCCAAAGGCTTTTACAAGCTTGCTGTCAGGAAAGATCAAAAGGTTTACCGGG 702
QY 300 IleAspAspProTyrGluProPheIleAsnGlyGluIleValIleLysMetLysAspGlu 319
DB 703 ATCGATGACCCCTTACGAGCCACCATGAACTCGAG-----ATTCTCTAGACGTGAA 756
QY 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrIleuGluAsnGly 339
DB 757 GGAGGAACCTTCCTTATCGAATGGCGGAAGAGGTCTCGGATACTTAGATAACAGGCT 816
QY 340 TyrLeuGlnAla 343
DB 817 TATCTTCAAGCA 828

RESULT 9

ADN74374

ID ADN74374 standard; cDNA; 831 BP.

XX ADN74374;
AC
XX
XX
DT 15-JUL-2004 (first entry)
XX
XX Thale cress cDNA repressed in E2Fa/Dpa expressing plants SeqID 2269.
DE
XX gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;
KW growth regulator; animal feed product; thale cress;
KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.
XX Arabidopsis thaliana.
XX WO2004035798-A2.
XX 29-APR-2004.
XX
XX 20-OCT-2003; 2003WO-EP011658.
XX
XX 18-OCT-2002; 2002EP-00079408.
XX (CROP-) CROPDESIGN NV.
XX
XX Inze D, De Veylder L, Vlieghe K;
PI
XX WPI; 2004-348466/32.
XX P-PSDB; ADN74375.
XX
XX Altering plant characteristics, useful for producing plants for enzyme or
PT pharmaceutical production comprises modifying in a plant, expression of
PT one or more nucleic acids and/or modifying level or activity of one or
XX more proteins.
PS Claim 1; SEQ ID NO 2269; 134pp; English.
XX
XX This invention relates to a novel method for altering one or more plant
CC characteristics. Specifically, it refers to identifying genes that are up
CC - or down-regulated in transgenic plants overexpressing the heterodimeric
CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
CC alter plant characteristics accordingly. The present invention describes
CC generating transgenic plants for the production of growth regulators,
CC enzymes, therapeutics, pharmaceuticals and animal feed products, where
CC the altered plant characteristics are selected from increased yield or
CC biomass, enhanced survival capacity, stress tolerance, plant architecture
CC or physiology, altered endoreplication, biochemistry, signal
CC transduction, storage lipid mobilisation and/or altered photosynthesis.
CC each relative to the corresponding wild type plants. Accordingly, these
CC sequences can also be useful as positive or negative selectable markers
CC during transformation of cells or tissues. The identified genes play a
CC role in a variety of biological processes such as DNA replication, cell
CC wall biosynthesis, nitrogen and/ or carbon metabolism or they function as
CC transcription factors. This polynucleotide sequence is thale cress cDNA
CC repressed 1.3 fold or more in plants overexpressing the E2Fa/Dpa
CC transcription factor, given in an exemplification of the invention.
XX
SQ Sequence 831 BP; 224 A; 137 C; 223 G; 247 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,42e-41 Length: 831
Score: 816.00 Matches: 159
Percent Similarity: 77.8% Conservative: 31
Best Local Similarity: 65.16% Mismatches: 44
Query Match: 44.69% Indels: 10
DB: 12 Gaps: 3

US-10-829-432-4 (1-343) x ADN74374 (1-831)

QY 100 GlyAlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgArgAsn 119
DB 127 GGATCTCAAACTCTGAGTCATAACAAAATGGATCTATTCTCGAGGTAAATCCATTAC 186
QY 120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139

```
Db 187 GGT-----CACACGGGACAAAAGCAAGGACCT---TTGTCTACG 222
QY 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyClnSerAspArgGln 159
Db 223 GTCGGAACTCGACAANTATAAGTGGCATGAATGTTCTGTGAGAAAGTTGATGACAG 282
QY 160 LysLeuLeuGlyClnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
Db 283 AGATTGCTTGATCAGAAAGGATGTGTGATTTGGTCCACCGTCTTAGTGTTCAGGGAAG 342
QY 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpVal 199
Db 343 AGTACTTGGCTTGCTTGAATCAGATGCTGTATCAAAAGGGAAGCTTTGTTATATT 402
QY 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
Db 403 CTTGATGGTGATTAAGTTAGGCATGGCTTAACCGTGATCTTAGCTTTAAAGCTGAGGAT 462
QY 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
Db 463 CGTGCAGAGAATATTCTGATAGCTGGAGAGTTGCTAAGCTTTTGGCGATGCTGAATA 522
QY 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
Db 523 ATCTGCATTGCCAGTTTGATATCTCTTATAGACAGATAGGACGCTTGTGAAAGTTTG 582
QY 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
Db 583 CTCCCGGAGGAGATTTGTTGAGGTGTCATGGATGATACCGCTTAGTGTTCGAGGCG 642
QY 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
Db 643 AGGGATCAAAGGGTCTTTACAAGCTTGCTCGTCAGGAAAGATCAAAAGGTTTACCGGG 702
QY 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
Db 703 ATCGATGACCCCTTACGAGCCACCATTTGAACCTCGAG-----ATTTCTCTAGGACGTGAA 756
QY 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
Db 757 GGAGGAACCTTCTTCATCGAATGCGGAAAGGTCGCGGATACTTTAGATACAAGGGT 816
QY 340 TyrLeuGlnAla 343
Db 817 TATCTTCAAGCA 828
RESULT 10
AAC35266
ID AAC35266 standard; DNA; 1068 BP.
XX
AC AAC35266;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 9570.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126788P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142390P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
```

PR	19-OCT-1999	99US-0144325P.	PR	12-OCT-1999	99US-0158369P.
PR	19-JUL-1999	99US-0144331P.	PR	13-OCT-1999	99US-0159299P.
PR	19-JUL-1999	99US-0144332P.	PR	13-OCT-1999	99US-0159294P.
PR	19-JUL-1999	99US-0144333P.	PR	13-OCT-1999	99US-0159295P.
PR	19-JUL-1999	99US-0144334P.	PR	14-OCT-1999	99US-0159329P.
PR	19-JUL-1999	99US-0144335P.	PR	14-OCT-1999	99US-0159330P.
PR	20-JUL-1999	99US-0144352P.	PR	14-OCT-1999	99US-0159331P.
PR	20-JUL-1999	99US-0144632P.	PR	14-OCT-1999	99US-0159637P.
PR	20-JUL-1999	99US-0144884P.	PR	14-OCT-1999	99US-0159638P.
PR	21-JUL-1999	99US-0144814P.	PR	18-OCT-1999	99US-0159584P.
PR	21-JUL-1999	99US-0145086P.	PR	21-OCT-1999	99US-0160741P.
PR	21-JUL-1999	99US-0145088P.	PR	21-OCT-1999	99US-0160767P.
PR	21-JUL-1999	99US-0145085P.	PR	21-OCT-1999	99US-0160768P.
PR	22-JUL-1999	99US-0145087P.	PR	21-OCT-1999	99US-0160770P.
PR	22-JUL-1999	99US-0145089P.	PR	21-OCT-1999	99US-0160814P.
PR	22-JUL-1999	99US-0145192P.	PR	21-OCT-1999	99US-0160815P.
PR	23-JUL-1999	99US-0145145P.	PR	22-OCT-1999	99US-0160980P.
PR	23-JUL-1999	99US-0145218P.	PR	22-OCT-1999	99US-0160981P.
PR	23-JUL-1999	99US-0145224P.	PR	22-OCT-1999	99US-0160989P.
PR	26-JUL-1999	99US-0145276P.	PR	25-OCT-1999	99US-0161404P.
PR	27-JUL-1999	99US-0145913P.	PR	25-OCT-1999	99US-0161405P.
PR	27-JUL-1999	99US-0145918P.	PR	25-OCT-1999	99US-0161406P.
PR	27-JUL-1999	99US-0145919P.	PR	25-OCT-1999	99US-0161359P.
PR	28-JUL-1999	99US-0145951P.	PR	26-OCT-1999	99US-0161360P.
PR	02-AUG-1999	99US-0146386P.	PR	26-OCT-1999	99US-0161361P.
PR	02-AUG-1999	99US-0146388P.	PR	28-OCT-1999	99US-0161920P.
PR	02-AUG-1999	99US-0146389P.	PR	28-OCT-1999	99US-0161992P.
PR	03-AUG-1999	99US-0147038P.	PR	28-OCT-1999	99US-0161993P.
PR	04-AUG-1999	99US-0147204P.	PR	29-OCT-1999	99US-0162142P.
PR	04-AUG-1999	99US-0147302P.	Alignment Scores:		
PR	05-AUG-1999	99US-0147192P.	Pred. No.:	1.87e-41	Length: 1068
PR	05-AUG-1999	99US-0147260P.	Score:	816.00	Matches: 159
PR	06-AUG-1999	99US-0147303P.	Percent Similarity:	77.8%	Conservative: 31
PR	06-AUG-1999	99US-0147416P.	Best Local Similarity:	65.16%	Mismatches: 44
PR	09-AUG-1999	99US-0147493P.	Query Match:	44.69%	Indels: 10
PR	09-AUG-1999	99US-0147935P.	DB:	3	Gaps: 3
PR	10-AUG-1999	99US-0148171P.	US-10-829-432-4 (1-343) x AAC35266 (1-1068)		
PR	11-AUG-1999	99US-0148319P.	Qy	100	GlyAlaArgThrHisCysHisargGlylleGlyArgGtrpValArgArgArgArgAsn 119
PR	12-AUG-1999	99US-0148341P.	Db	220	GGATCTCAAACTCTGAGTCATAACAAAATGGATCTATTCTCGAGGTTCATTCATTCATTAAC 279
PR	13-AUG-1999	99US-0148565P.	Qy	120	GlyAlaAlaProGlyGluAlaProHisSerProValHisGlySerProValMetSerAsn 139
PR	13-AUG-1999	99US-0148684P.	Db	280	GGT-----CACACGGGCAAAAGCAGGACCT---TTGTCTACG 315
PR	16-AUG-1999	99US-0149368P.	Qy	140	IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
PR	17-AUG-1999	99US-0149175P.	Db	316	GTGGAAACTCGACAAATATAAGTGGCATGAATGTTCTGTTCAGAAAAGTTGATAGACAG 375
PR	18-AUG-1999	99US-0149426P.	Qy	160	LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
PR	20-AUG-1999	99US-0149723P.	Db	376	AGATTGCTTGATCAGAAAGAGTGTGATTGGGTTCACCGGTCTTAGTGGTTCAGGGAG 435
PR	20-AUG-1999	99US-0149829P.	Qy	180	SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpVal 199
PR	23-AUG-1999	99US-0149902P.	Db	436	AGTACTTTGGCTTGTGCTTTGAATCAGATGTTGTATCAAAAGGGAAGCTTTGTTATATT 495
PR	23-AUG-1999	99US-0149930P.	Qy	200	LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheIysAlaGluAsp 219
PR	25-AUG-1999	99US-0150566P.	Db	496	CTTGATGGTGATAATGTTAGGCATGCTTTAAACCGTGATCTTAGCTTTAAAGCTCAGGAT 555
PR	26-AUG-1999	99US-0150884P.	Qy	220	ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239


```
Db      676  CTCCCCGAGGAGATTTGTTGAGGTGTTTCATGATGTACCGCTTAGTGTTCGCGAGGG 735
Qy      280  AtgAspProlysglyLeuTyrlyLeuAlaAargThrGlyVallelysglypethrGly 299
Db      736  AGGATCCAAAGGGTCTTTACAGCTTGCTCGTCAGGAAGATCAAGGTTTACCCGG 795
Qy      300  ileAspAspProTyrGluProIleAsnGlyGluIleVallelysglyMetlyAspGlu 319
Db      796  ATCGATGACCCITACGAGCCACCATTGAATCGGAG-----ATTCTCTAGGACGTGAA 849
Qy      320  GluCysProSeuProlysalalMetAlaLysGlnValLeuCyTyrLeuGluGluAenGly 339
Db      850  GGAGGAACCTTCTCTATCGAAATGCGGAAAGGTGCTCGGATACTTAGATAACAAGGCT 909
Qy      340  TyrLeuGlnAla 343
Db      910  TATCTTCAAGCA 921

RESULT 11
AAC45659
ID AAC45659 standard; DNA; 1183 BP.
XX
AC AAC45659;
XX
XX 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 47294.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 19-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
```

```
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.

PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.

PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 2.08e-41 Length: 1183
Score: 816.00 Matches: 159
Percent Similarity: 77.87% Conservative: 31
Best Local Similarity: 65.16% Mismatches: 44
Query Match: 44.69% Indels: 10
DB: 3 Gaps: 3

US-10-829-432-4 (1-343) x AAC45659 (1-1183)
QY 100 GlyAlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgAsn 119
DB 221 GGATCTCAACTCTGAGTCATAACAAAATGATCTATCTCTGAGTTAAATCCATTAC 280
QY 120 GlyAlaAlaProGlyGluAlaProHisSerProValysGluLysProValMetSerAsn 139
DB 281 GGT-----CACACGGGACAAAAGCAAGGACCT---TTGTCTAGC 316
QY 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
DB 317 CTCGAAACTCGACAAATATAAAGTGGCATGAATGTTCTGTGAGAAAAGTTGATAGACAG 376
QY 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
DB 377 AGATTGCTTGATCAGAAAGGATGTGATTTGGGTACCCTCTTAGTGGTTCAGGGAAG 436
QY 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
DB 437 AGTACTTTGGCTTGCTTTGAATCAGATGTTTATCAAAAGGGGAAGCTTTGTATATAT 496
QY 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
DB 497 CTTGATGTTGATAATGTTAGGCATGGCTTAAACCGTGATCTTAGCTTTTAAAGCTGAGGAT 556
QY 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
DB 557 CGTCAGAGATATTCGTAGAGTTGGAGAGTTGCTTAGCTTTTTCGGATGCTGGAATA 616
QY 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
DB 617 ATCTGCATTGGCAGTTTGATATCTCTTATAGAACAGATAGGACGCTTGCGAAGTTTG 676
QY 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
DB 677 CTCCTCCGAGGAGATTTTGTGAGGTTCATGATGATGATGATGATGATGATGATGATGATG 736
QY 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
DB 737 AGGATCCAAAGGGTCTTTACAGCTTCTGTCAGGAAAGATCAAAAGGTTTACCGGG 796
QY 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
DB 797 ATCGATGACCTTACGAGCCACCATTAACATGCGAG-----ATTTCCTAGGACGTGAA 850
QY 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
DB 851 GGAGGAACCTTCTCTATCGAAATGCGGAAAGGTCGTCGGATACTTAGATAACAAGGCT 910
QY 340 TyrLeuGlnAla 343
DB 911 TATCTTCAAGCA 922

RESULT 12
AAC51190
ID AAC51190 standard; DNA; 714 BP.
XX
AC AAC51190;
XX
DT 18-OCT-2000 (first entry)
XX
```


US-10-829-432-4 (1-343) x AAC51190 (1-714)

Qy	137 MetSerAsnIleGlyLyssSerThrAsnIleLeuTrpHisAanCysLeuIleGlyGlnSer	156 :::
Db	1 AUTCACAGTGGGAATTCAACGACATATTTGGCAAGAATCCCCCATTTGGAAACT	60 :::
Qy	157 AspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGly	176 :::
Db	61 GAAGGCAGCAGCTTGTAATAACAGAACGGTGTGTGTGTGTCATCACAGCGCTCAGTGGC	120 :::
Qy	177 SerGlyLYsaSerThrlEuAlaCySaAlalaSerArgGluLeuHIScYsaRGlyHiSLeu	196 :::
Db	121 TCAGGGAAAAGCAGCTTAGCTTCTCGCTTAGTAGAGAGCTGAACAACCCGGGAAACCTA	180 :::
Qy	197 ThrTyrrValLeuAspGlyAspAsnLeuArghisGlyfLeuAsnArgraspLeuSerPheLys	216 :::

(ZHUT/) ZHU T.

Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
Goff SA, Katagiri F, Kreps J, Provart N, Rieke D, Zhu T;
WPI; 2004-190374/18.

New rice promoter, useful for manipulating crop plants to alter or
improve phenotypic characteristics, e.g. produce large quantities of oil
or proteins, resistance to insecticides, virus or fungi, stress tolerance
or high nutritional value.

Claim 72; SEQ ID NO 869; 230pp; English.

The invention relates to plant nucleotide sequences that direct seed-,
leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
or constitutive transcription of an operatively linked nucleic acid
segment. The invention also relates to a method for augmenting a plant
genome and a method of identifying a gene, where its expression is
altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
sorghum, rice or wheat. The polynucleotides and the polypeptides they
encode are useful for manipulating crop plants to alter or improve
phenotypic characteristics, to produce large quantities of oil or
proteins, to incur resistance to insecticides, viruses or fungi, and to
incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
have a high nutritional value with reduced apical dominance or dwarfism,
early flowering or altered metabolic pathways. This sequence represents a
plant nucleic acid of the invention. Note: The sequence data for this
patent did not form part of the printed specification but was obtained in
electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

Sequence 687 BP; 195 A; 154 C; 165 G; 173 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6,11e-40 Length: 687
Score: 788.00 Matches: 158
Percent Similarity: 83.89% Conservativeness: 19
Best Local Similarity: 74.88% Mismatches: 21
Query Match: 43.15% Indels: 13
DB: 12 Gaps: 2

US-10-829-432-4 (1-343) x ADJ39869 (1-687)

QY 138 SerAenlleGlyLysSerThrAenlleLeuTrpHisAenCysLeuIleGlyClnSerAep 157
DB 10 TCAACTGTCGGTAAATCAACAACATCCGATGCGATGACTGCCAGTAAACAGCTTGAC 69

QY 158 ArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGly--- 176
DB 70 AGACAAAGTCTGCTACCAAGAGGGCTGTGTGTTGGATCACTGGCTTAAGTGTTCA 129

QY 177 -----SerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgG1 194
DB 130 GTCAACTTCAGGAAAAAGCACCTTGCATGTGCTACGTGACCGCTGAGCTCAAGAGG 189

QY 194 yHisLeuThrTrpValLeuAspGlyAspAenLeuArgHisGlyLeuAenArgAspLeuSe 214
DB 190 GCATCTGACCTTATGTTCTTGACGGGACAAATCTCCGGCAGCGCTGAAACAAGATCTCAG 249

QY 214 rPhelyAlaGluAspArgAlaGluAenlleArgArgVal-GlyGluValAlaLysLeuP 234
DB 250 CTTCAAGGCAAGAGTCTGCGAANAATATACGAGAGTTGGAGAGTGGCAAGCTGT 309

QY 234 heAlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrrArgArgAspArgA 254
DB 310 TTGCAGATGCTGGATTGATCTGCATTACTTACTAGTTGATATCACTCCCTATAAGAGTGATCGAA 369

QY 254 sPAlaCysArgAlaLeuLeuProHisSerAenPheileGluValPheleAspLeuProL 274
DB 370 GCGCGCTGCCGCAAAATTATACCAATCTTCGTTTCAATGAGGTGTTCTGTAAGTGTCCAC 429

QY 274 euLysIleCysGluAlaAlaArgAspProLysGlyLeuTyrrLysLeuAlaAargThrGlyLysI 294
DB 430 TTGAAGTATGTGACAGAAAGGGATCCAAAGGCTGTGACAGCTTCTGTCGCGCAAAA 489

QY 294 le-LysGlyPheThrGlyIleAspAspProTyrrGluProProlleAsnGlyGluIleVal 313
DB 490 TCAAGGGCTTTACGGGAATAGATGATCTTATGAAACACCTTCAGATTGTGAG----- 543

QY 314 lIleLysMetLysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCys 333
DB 544 -----GGGAGTGGCTTCACCTAANTCAATGGCTGATCAAGTAGTGTC 588

QY 334 TyrLeuGluGluAenGlyTyrrLeuGln 342
DB 589 TATCTTGAAGCAATGGATTCTTTTCAG 615

RESULT 14
AAZ98337
ID AAZ98337 standard; DNA; 917 BP.
XX
AC AAZ98337;
XX
DT 14-JUN-2000 (first entry)
XX
DE A. thaliana gene involved in environmental stress tolerance.
XX
KW Environmental stress; plant; transgenic plant; anaerobic; flooding; cold;
KW dehydration; drought; heat stress; salinity; osmotolerance; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200008187-A2.
XX
PD 17-FEB-2000.
XX
PF 04-AUG-1999; 99WO-EP005652.
XX
PR 04-AUG-1998; 98EP-00202634.
XX
PA (VLAA-) VLAAIS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI Lee JH, Verbruggen N;
XX
DR WPI; 2000-205726/18.
XX
PT P-PSDB: AAY77957.
XX
PS Isolation of polynucleic acids useful for producing transgenic plant by
isolating genes involved in tolerance to environmental stress.
XX
PS Claim 4; Page 218-220; 312pp; English.
XX
CC The invention relates to isolation of coding sequences and/or genes
involved in tolerance to environmental stress in plants. The sequences
(AAZ98305-298365) are useful for producing a transgenic plant having
enhanced tolerance or resistance to environmental stress conditions such
as anaerobic, flooding, cold, dehydration, drought, heat stress or
salinity. This is useful for producing improved yield, growth,
development and productivity under environmental stress conditions, and
also provides growth of crops in areas where they cannot grow without the
induced osmotolerance. Sequences AAZ98305-365 represent polynucleotide
sequences from A. thaliana that are involved in environmental stress
tolerance
XX
SQ Sequence 917 BP; 244 A; 220 C; 229 G; 224 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.21e-39 Length: 917
Score: 778.50 Matches: 175
Percent Similarity: 66.04% Conservativeness: 37
Best Local Similarity: 54.52% Mismatches: 67
Query Match: 42.63% Indels: 43
DB: 3 Gaps: 9

US-10-829-432-4 (1-343) X AAZ98337 (1-917)

Qy	40	ProThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaProProVal-----57
Db	14	CCGTCGGATTCAATGGAAGATTAGCTATCAGACGATCTCGA-----CGTCGGTTTTCTGT70
Qy	58	---LeuProGlyLeu-----ThrProSerAspAla-----Pro67
Db	71	TCATATTCAGAGTCTCGCGCGCGAATCCACCGAATACTCCAAGTGACGGTTTCTCTCAAG130
Qy	68	LeuProAlaLeuValIleHisGlyLeuThrProArgSerSerHisSerSerAlaGlyLeu87
Db	131	CTGCTTCGCTC-GTCTATTCCGCGCAGACCGGAAATAATTAGTAGCGAATTCTACTTCCITT189
Qy	88	AlaSerAspSerGlyArgArgGluGlyArgGlyAlaAargThrHisCysHisAsArg107
Db	190	TCATCAAACTCTCAGCCGT-----TAAAGTCTCTCTCCTCAAGCTTCCCTCAACCGC237
Qy	108	GlyIleGlyArg-----TrpValArgArg-ArgArgArgAsnGlyAlaAlaPr123
Db	238	TGATTTTCCGCGCTTTCAGAACTATACTGAAAGAGGGAAGAAATAACCGA-----289
Qy	123	oGlyGluAlaProHisSerProValLysGluLysProValMetSerAsnIleGlyLysSe143
Db	290	-----AAAGAG-----AAAGC300
Qy	143	rThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuGlu163
Db	301	AGAGAATCATCGTGTGCGACGAGATTGCATATGCAGATGCGACAGACAACAATCTTCA360
Qy	163	yGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAl183
Db	361	ACAAAGGGTCTGTGCTTGGATCCTCTCAGTGTCTCAGTGTTCAGGGAAGAAGCACTGTGC420
Qy	183	aCysAlaLeuSerArgGluLeuHisCysAargGlyHisLeuThrTyrValLeuAspGlyAs203
Db	421	TTGTGCACCTAAGTAAAGCACTTTTCAAGAGGCGAAACTTACTTACACACTCGACGCGCA480
Qy	203	pAsnLeuArgHisGlyLeuAsnArgAspLeuSer-PhelysAlaGluAspArgAlaGluAs223
Db	481	CAATGTCGTACGGCCCTTAAACCGTGACCTCCTTCTCAAGCAGACGACCGACCGGAAA540
Qy	223	nIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAl243
Db	541	CATTAGAAGAAATTTGGTGAGTGGCTAAGTTGTTTCTGACGTCGAGCTCAATTGTATAGC600
Qy	243	aSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeuLeuProHisSe263
Db	601	AACTTTTGATTTCTCGATCCGAGAGACAGACGCGTCCGCTCTTGTGTACCTGACGCG680
Qy	263	rAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaAargAspProLy283
Db	661	CGATTTTCGTGAGGCTTTCATGCGAGTTCCTCTTCATGTGTGCGAGTCGAGATCCNAA720
Qy	283	sGlyLeuTyrLysLeuAlaAargThrGlyLysIleLysGlyPheThrGlyIleAspAspPr303
Db	721	GGGGTTGTACAAGCTTGCAGTGCAGGCGAATAAAGGCTTCACTGGAATCGACGACCC780
Qy	303	oTyrGluProProIleAsnGlyGluIleValIleLys---MetLysAspGluGluCysPr322
Db	781	TTACAGAGGCACCATGTAATTCGAGGGTAGTGTCTGAACACACACAGGACACCGACGTCGT840
Qy	322	oSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGlyTyrLeuGlu342
Db	841	TTCCGACAGTCAGATGGCTGAGAACATCATCTCTTACCTCGAANAACAAAGGTTATCTTGA900
Qy	342	n342
Db	901	G901

RESULT 15

AAC42294

ID AAC42294 standard; DNA; 948 BP.

XX	AAC42294;
XX	AC
XX	AD
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana DNA fragm
XX	
KW	Hybridisation assay; genetic
KW	protein identification; signal
KW	promoter; termination sequen
XX	
OS	Arabidopsis thaliana.
XX	
PN	EPI033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-00301439
XX	
PR	25-FEB-1999;
PR	99US-0121825P
PR	05-MAR-1999;
PR	99US-0123380P
PR	23-MAR-1999;
PR	99US-0125788P
PR	25-MAR-1999;
PR	99US-0126264P
PR	29-MAR-1999;
PR	99US-0126785P
PR	01-APR-1999;
PR	99US-0127462P
PR	06-APR-1999;
PR	99US-0128234P
PR	08-APR-1999;
PR	99US-0128714P
PR	16-APR-1999;
PR	99US-0129845P
PR	19-APR-1999;
PR	99US-0130077P
PR	21-APR-1999;
PR	99US-0130449P
PR	23-APR-1999;
PR	99US-0130510P
PR	25-APR-1999;
PR	99US-0130891P
PR	28-APR-1999;
PR	99US-0131449P
PR	30-APR-1999;
PR	99US-0132048P
PR	03-MAY-1999;
PR	99US-0132407P
PR	04-MAY-1999;
PR	99US-0132484P
PR	05-MAY-1999;
PR	99US-0132485P
PR	06-MAY-1999;
PR	99US-0132486P
PR	07-MAY-1999;
PR	99US-0132487P
PR	11-MAY-1999;
PR	99US-0132863P
PR	12-MAY-1999;
PR	99US-0134256P
PR	14-MAY-1999;
PR	99US-0134218P
PR	14-MAY-1999;
PR	99US-0134219P
PR	14-MAY-1999;
PR	99US-0134321P
PR	14-MAY-1999;
PR	99US-0134370P
PR	18-MAY-1999;
PR	99US-01343768P
PR	19-MAY-1999;
PR	99US-0134941P
PR	20-MAY-1999;
PR	99US-0135242P
PR	21-MAY-1999;
PR	99US-0135353P
PR	24-MAY-1999;
PR	99US-0136029P
PR	25-MAY-1999;
PR	99US-0136021P
PR	27-MAY-1999;
PR	99US-0136392P
PR	28-MAY-1999;
PR	99US-0136782P
PR	01-JUN-1999;
PR	99US-0137222P
PR	03-JUN-1999;
PR	99US-0137528P
PR	04-JUN-1999;
PR	99US-0137502P
PR	07-JUN-1999;
PR	99US-0137724P
PR	08-JUN-1999;
PR	99US-0138094P
PR	10-JUN-1999;
PR	99US-0138547P
PR	10-JUN-1999;
PR	99US-0138847P
PR	14-JUN-1999;
PR	99US-0139119P
PR	16-JUN-1999;
PR	99US-0139452P
PR	16-JUN-1999;
PR	99US-0139453P
PR	17-JUN-1999;
PR	99US-0139492P
PR	17-JUN-1999;
PR	99US-0139454P
PR	18-JUN-1999;
PR	99US-0139455P
PR	18-JUN-1999;
PR	99US-0139456P
PR	18-JUN-1999;
PR	99US-0139457P
PR	18-JUN-1999;
PR	99US-0139458P
PR	18-JUN-1999;
PR	99US-0139459P
PR	18-JUN-1999;
PR	99US-0139460P
PR	18-JUN-1999;
PR	99US-0139461P

[illegible]

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 9, 2005, 13:05:12 ; Search time 198 Seconds
(without alignments)
2834.560 Million cell updates/sec

Title: US-10-829-432-4
Perfect score: 1826
Sequence: 1 RPFHINOTEPVLTHTQOPP.....PRMAKQVLCYLENGYLQA 343

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp
-Q=/cgn2_1/USFTO_spool_p/US10829432/runat_08092005_132224_25102/app_query.fasta_1.519
-DB=Issued Patents NA -OFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOFC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10829432@cgn.1.105@runat_08092005_132224_25102 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	521.5	28.6	615	US-09-134-001C-1680	Sequence 1680, Ap
2	517	28.3	681	US-09-543-681A-2995	Sequence 2995, Ap
3	503	27.5	651	US-09-489-039A-2803	Sequence 2803, Ap
4	496	27.2	2014	US-09-898-165B-1	Sequence 1, Appli
5	495	27.2	2506	US-08-879-561-4	Sequence 4, Appli
6	495	27.1	1875	US-09-949-016-1468	Sequence 1468, Ap
7	494.5	27.1	1875	US-09-949-016-1925	Sequence 1925, Ap
8	486.5	26.6	2617	US-09-786-240-21	Sequence 21, Appl
9	484.5	26.5	1851	US-09-898-165B-10	Sequence 10, Appl
10	484.5	26.5	2000	US-09-898-165B-2	Sequence 2, Appli
11	482	26.4	1845	US-09-898-165B-9	Sequence 9, Appli
12	475	26.0	1160	US-09-153-310-41	Sequence 41, Appl

c	13	458.5	25.1	640681	4	US-09-790-988-1	Sequence 1, Appli
	14	402.5	22.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	15	402.5	22.0	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	16	377.5	20.7	26289	4	US-09-902-540-1210	Sequence 1210, Ap
	17	339	18.6	546	4	US-09-902-540-4227	Sequence 4227, Ap
	18	318.5	17.4	1959	4	US-09-252-991A-7983	Sequence 7983, Ap
c	19	318.5	17.4	2010	4	US-09-252-991A-8169	Sequence 8169, Ap
c	20	261.5	14.3	601	4	US-09-949-016-50640	Sequence 50640, A
	21	261.5	14.3	109925	4	US-09-949-016-13210	Sequence 13210, A
	22	251.5	13.8	759	4	US-09-902-540-8116	Sequence 8116, Ap
c	23	251.5	13.8	11502	4	US-09-902-540-1064	Sequence 1064, Ap
	24	245	13.4	89892	4	US-09-949-016-13667	Sequence 13667, A
	25	184	10.1	53500	4	US-09-266-965-76	Sequence 76, Appl
c	26	156	8.5	9937	4	US-09-902-540-946	Sequence 946, App
	27	155	8.5	6803	3	US-08-665-259-19	Sequence 19, Appl
	28	155	8.5	6803	3	US-08-762-500-19	Sequence 69, Appl
	29	150.5	8.2	483	4	US-09-266-965-69	Sequence 24, Appl
	30	146.5	8.0	2323	3	US-09-149-476-24	Sequence 8948, Ap
c	31	138.5	7.6	7680	4	US-09-902-540-8948	Sequence 955, App
	32	138.5	7.6	11476	4	US-09-902-540-955	Sequence 11768, A
	33	137.5	7.5	1479	4	US-09-252-991A-11768	Sequence 11840, A
	34	137.5	7.5	1968	4	US-09-252-991A-11840	Sequence 1207, Ap
	35	132.5	7.3	23417	4	US-09-902-540-1207	Sequence 5422, Ap
	36	131.5	7.2	2130	4	US-09-252-991A-5422	Sequence 17448, A
	37	127	7.0	58014	4	US-09-949-016-17448	Sequence 709, App
c	38	126.5	6.9	4183	4	US-09-902-540-709	Sequence 13724, A
	39	125	6.8	30782	4	US-09-949-016-13724	Sequence 10, Appl
c	40	125	6.8	54945	4	US-09-967-669-10	Sequence 17187, A
	41	124.5	6.8	10106	4	US-09-949-016-17187	Sequence 12369, A
	42	124.5	6.8	27933	4	US-09-949-016-12369	Sequence 1033, Ap
	43	124	6.8	3446	4	US-09-949-016-1033	Sequence 5776, Ap
	44	124	6.8	3446	4	US-09-949-016-5776	Sequence 5777, Ap
	45	124	6.8	3446	4	US-09-949-016-5777	

ALIGNMENTS

RESULT 1

US-09-134-001C-1680

; Sequence 1680, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 1680

; LENGTH: 615

; TYPE: DNA

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-1680

Alignment Scores:					
Pred. No.:	4.17e-36	Length:	615		
Score:	521.50	Matches:	102		
Percent Similarity:	67.33%	Conservative:	34		
Best Local Similarity:	50.50%	Mismatches:	63		
Query Match:	28.56%	Indels:	3		
DB:	3	Gaps:	1		

US-10-829-432-4 (1-343) x US-09-134-001C-1680 (1-615)

Qy 140 lleGlyLysSerThrAsnIleuTrpHisAnCysLeuIleGlyGlnSerAspArgGln 159

Db 16 ATGAGTGAATCAATCATATTACATGGCATGACTCAGAAAGTTACGAGAAACAAAGACAG 75

QY 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
Db 76 CATAAAAATGGACACAAAAGTCTGTATATATGTTTACAGGGCTATCTGGGTGAGGAAA 135
QY 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
Db 136 TCAACGGTCTCTGTTGGCTTAGAAAAGAAATTTATTAATGAAGGAAAAACAAACATACCGT 195
QY 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
Db 196 TTAGATGGTGAATGTGGTCTATGCTTAATAAAAAATTTAGGGTTAGTCTCGAAGAT 255
QY 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
Db 256 CGTTCAGAAAATATTCGACGCATTGGAGAAGTAGCAAAATTAATGGTAGACCTGGAGCC 315
QY 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
Db 316 TTAACAGTACTGCTTTATCTCCCATATAAAGAGACAGAGAAGGTGTAGAGCATTA 375
QY 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
Db 376 CTAGAGGATATGAGTTTATAGAGTATATACAAAATGTAGTGTGGGAATGTGAAAAG 435
QY 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
Db 436 AGAGATCTCTAAAGATTGTATAAAAAGCAGCATCTGGGGAAATACCTGAAATTTACAGT 495
QY 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
Db 496 ATAAGTGCACTTATCAACACCTGAAAACCCCTGAAATTTACTATA-----GATACT 546
QY 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
Db 547 GAACAGCACATTTGAACAATCAGTGGTACAAAATTTACGTTTCTTTAAGAACACGAA 606
QY 340 TyrLeu 341
Db 607 TATATT 612

RESULT 2

US-09-543-681A-2995
; Sequence 2995, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 2995
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-2995

Alignment Scores:
Pred. No.: 1.19e-35 Length: 681
Score: 517.00 Matches: 107
Percent Similarity: 60.79% Conservative: 31
Best Local Similarity: 47.14% Mismatches: 57
Query Match: 28.31% Indels: 32
DB: 4 Gaps: 4

US-10-829-432-4 (1-343) x US-09-543-681A-2995 (1-681)

QY 133 GluLysProValMetSerAsnIleGlyLysSer----- 143
Db 13 GATCCGCCGCCATTTTCCCGCATTTGGGGCCCAAGAGACTTACTGGGAGGAAATAGTGAGG 72

QY 144 -----ThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeu 161
Db 73 ATACATCAAGATATGTCTGGCATCCTCATCAATAGGTTAAAAAGCGTGAACACAA 132
QY 162 LeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerThr 181
Db 133 CAGGTACACAAGAGTGTACTTGTGGTTTACTGGGTATCTGGGTACAGTAAATCAACA 192
QY 182 LeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeu----- 196
Db 193 CTGGTGTATGGCTAGACCAACCTTATATCAGTACTCGACACTCCATCGCCCTATTCGC 252
QY 197 ThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLys 216
Db 253 ACCTATTATTAGATGGTGATATCTACGCCATGGTTTATGCCATGATCTTTGGGTTTAGT 312
QY 217 AlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAsp 236
Db 313 GAACAAGATAGGCATGAAATATTTCGGCTGTAGGGAAATGGCTAAATTAATGGTTGAT 372
QY 237 AlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCys 256
Db 373 GCCGGATTAAATGTCTTAAACAGCATTTATTTCTCTATCAGCAAGATAGACAAAGTA 432
QY 257 ArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIle 276
Db 433 AGAGAAAGTTTGGTCAAGGGCGATTATTATTCAGATCTTTGTTGATACACCTTTAGCCCTT 492
QY 277 CysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGly 296
Db 493 TGTGAAGCACGTGATCTCTAAAGGCTCTATCAAAAAGCGCAGGAGGAGATCAAAACAG 552
QY 297 PheThrGlyIleAspAspProTyrGluProPro-----IleAsnGly 310
Db 553 TTTTCGGCATTTGATTCACCTTATGAACCCCCCTCGCCGAGAAATTCATTATAGCCGC 612
QY 311 GluIleValIleLysMetLysAspGluGluCysProSerProLysAlaMetAlaLysGln 330
Db 613 TCACCTCGCGATT-----AATGACCTTACACAA 642

RESULT 3

US-09-489-039A-2803
; Sequence 2803, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2803
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2803

Alignment Scores:
Pred. No.: 1.79e-34 Length: 651
Score: 503.00 Matches: 97
Percent Similarity: 70.24% Conservative: 21
Best Local Similarity: 57.74% Mismatches: 50
Query Match: 27.55% Indels: 0
DB: 4 Gaps: 4

US-10-829-432-4 (1-343) X US-09-489-039A-2803 (1-651)

[illegible]

RESULT 4

```

US-09-898-165B-1
; Sequence 1, Application US/09898165B
; Patent No. 6818428
; GENERAL INFORMATION:
; APPLICANT: Daniel H. Cohn
; APPLICANT: Muhammad Faiyaz ul Haque
; APPLICANT: Lily M. King
; APPLICANT: Deborah Krakow
; TITLE OF INVENTION: 3-Phosphoadenosine-5-Phosphosulfate
; TITLE OF INVENTION: (PAPS) Synthetase Proteins and Methods for Treating
; TITLE OF INVENTION: Osteoarthritic Disorders
; FILE REFERENCE: 18810-81552
; CURRENT APPLICATION NUMBER: US/09/898,165B
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/399,212
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; LENGTH: 2014
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-898-165B-1

```

Alignment Scores:		
Pred. No.:	3.81e-33	2014
Score:	496.00	Matches: 108
Percent Similarity:	62.73%	Conservative: 30
Best Local Similarity:	49.09%	Mismatches: 61
Query Match:	27.16%	Indels: 21
DB:	4	Gaps: 5

US-10-829-432-4 (1-343) x US-09-898-165B-1 (1-2014)

114	Qy	ArgArgArgArgArg	-----AsnClyAlaalaProGly-GluAlaProHis	121
11	Db	CGCGCGCGCGCGCT	CTCTGGTCTCTGCTCCGGGACCGCGCTCCCGCGA	70
129	Qy	rProVallys	-----GluLysProValMetSerAsnIleGlyLysSerThrAs	145
71	Db	GCCAGCCAGCATGTCGGGGAT	CAAGAAAGACGAGAGAACAGCAGCAAGAAATCCACCAA	130
145	Qy	nIleLeuThrPhiAsnCysLeuIleGlyGlnSerAspArgGlnIlysLeuLeuGlyGln--	164	
131	Db	TGTAGTCTATCAGGCCCCACCATG	TGTAGCAGGAATAAGAGAGGGCAAGTGGTTGSAACAAAG	190
165	Qy	-----LysGlyCysValValTrrIleThrGlyLeuSerGlySerGlyLysSerTh	181	
191	Db	GGGTGGGTTCCGAGGATGTC	CGTGTGCTAACAGGCTCTCTGTGCTGGGAAACAC	250
181	Qy	rLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAs	201	
251	Db	GATAAGTTTTGCCCTGGAGGAGTACCT	TGTCTCCATGCCATCCCTGTTACTCCTCGA	310
201	Qy	pGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAl	221	
311	Db	TGGGGCAATGTCTGTCATGCGCTTAA	CAGAAATCTCGATCTCTCTCGGGGACAGAGA	370
221	Qy	aGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaaspAlaGlyValIleCy	241	
371	Db	GGAAATATCCGCGGATGCTGAGGTGGCTTA	GCTGTTGCTGATGCTGGTCTGGTCTG	430
241	Qy	sIleAlaSerLeuIleSerProTyrArgArgAspAlaCysArgAlaLeu-----	259	
431	Db	CATTACCAGCTTTTATTTCTCCATTCGCAAGAGATCGT	GAGAAATGCCCGCAAAATACATGA	490
260	Qy	-----LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIle	276	
491	Db	ATCAGCAGGGCTGCGCA-----TTCTTTTGAATATTTGTAGATGCACCTCTAAATAT	541	
276	Qy	eCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysG	296	
542	Db	TTGTGAAGCAGACGCTAAAGCCCTCTATAAAGGCCACAGACTGGGGAGATTAAAGG	601	
296	Qy	yPheThrGlyIleAspAspProTyrGluProProIleAsnGlyGluIleValIleLys	315	
602	Db	ATTTACAGGTATTGATTCTGATTATGAGAAACCTGAAACTCCTGAGCGTGTCTTAA	659	

RESIT 5

US-08-879-561-4
; Sequence 4, Application US/08879561
; Patent No. 5817482
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/879,561
; FILING DATE: Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

Db	331	GGCCTTAACAGAAATCTCGGAATCTCTCCTCGGGGACAGAGAGAAAAATATATCCGCGGATT	399
Qy	228	GlyGluValAlaIalysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSer	247
Db	391	GCTGAGGTGGCTAGACTGTTTGGCTGATGCTGGTCTGGCTTGCATTTACCAGCTTTATTTCT	450
Qy	248	ProTyrArgArgAspArgAspAlaCysArgAlaLeu-----LeuProHis	262
Db	451	CCATTTCGCAAGAGTTCGTGAGATGCTCCGCCAAAAATACATGAATCAGCAGGGCTGCCA----	507
Qy	263	SerAsnPheIleGluValPheIleAspLeuProlLeuLysIleCysGluAlaArgAspPro	282
Db	508	-----TTCCTTTGAAATATTTTGTAGATGCACCTCTAAATATTTTGTGAAGCAGAGACGTA	561
Qy	283	LysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAsp	302
Db	562	AAAGCCTCTATAAAGGCCAGAGCTGGGGAGATTAAGGATTTTACAGGTATTGATTCT	621
Qy	303	ProTyrGluProIleAsnGlyGluIleValIleLys	315
Db	622	GATTATGAGAAACCTGAAACTCCTGAGCGTGTGCTTAAA	660

RESULT 8
 US-09-786-240-21
 ; Sequence 21, Application US/09786240
 ; Patent No. 6558935
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE PHARMACEUTICALS, INC.
 ; APPLICANT: TANG, Y. Tom
 ; APPLICANT: CORLEY, Neil C.
 ; APPLICANT: GUEGLER, Karl J.
 ; APPLICANT: BAUGHN, Mariah R.
 ; APPLICANT: LAL, Preeti
 ; APPLICANT: YUE, Henry
 ; APPLICANT: HILLMAN, Jennifer L.
 ; APPLICANT: AZIMZAI, Yalda
 ; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
 ; FILE REFERENCE: PF-0592 PCT
 ; CURRENT APPLICATION NUMBER: US/09/786,240
 ; CURRENT FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: 09/150,657; unassigned; 09/186,779; unassigned; 6/
 ; PRIOR FILING DATE: 1998-09-10; 1998-09-10; 1998-11-04; 1998-11-04; 1999-05-
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 21
 ; LENGTH: 2617
 ; TYPE: DNA

[illegible]

Db 162 CATGTGAGCAGGAATAAGAGAGGGCAAGTGGTTGGAACAAGGGGTGGGTTCGAGGATGT 221
Qy 168 ValValTrpIleThrGlyLeuSerGlySerGlyLeuSerThrLeuAlaCysAlaLeuSer 187
Db 222 ACCGTGTGGCTAACAGGTCTCTCTGGTGTGGAACCAACCAATAGATTTTGGCCCTGGAG 281
Qy 188 ArgGluLeuHisCysArgGlyHisLeuThrTrpValLeuAspGlyAspAsnLeuArgHis 207
Db 282 GAGTACCTTGTCTCCATGCCATCCCTTGTACTCCCTGGATGGGCAATATGTCGTCTAT 341
Qy 208 GlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIleArgVal 227
Db 342 GCGCTTAACAGAAATCTCGATCTCTCTGGGGACAGAGAGAAAATATCCGCGGAT 401
Qy 228 GlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSer 247
Db 402 GCTGAGGTGGCTTAAGCTGTCTGATGCTGTCTGGTCTGCAATACAGCTTTATTTCT 461
Qy 248 ProTyrArgArgAspAlaCysArgAlaLeu-----LeuProHis 262
Db 462 CCATTCCGAAGGATCGTGAGATGCCCGCAAAATACATGATCAGAGGGCTGCCA--- 518
Qy 263 SerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspPro 282
Db 519 -----TTCTTTGAAATATTTGTAGATGCACCTCTAAATATTTGTGAAAGCAGACGTA 572
Qy 283 LysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAsp 302
Db 573 AAAGGCTCTATAAAGGCCAGAGCTGGGAGATTAAGGATTTACAGGATTAATGTTCT 632
Qy 303 ProTyrGluProPheIleAsnGlyGluIleValIleLys 315
Db 633 GATTATGAGAAACCTTGAAACTCTTGAGCGTGTGCTTAAA 671
RESULT 9
US-09-898-165B-10
; Sequence 10, Application US/09898165B
; Patent No. 6818428
; GENERAL INFORMATION:
; APPLICANT: Daniel H. Cohn
; APPLICANT: Muhammad Faiyaz ul Haque
; APPLICANT: Lily M. King
; APPLICANT: Deborah Krakow
; TITLE OF INVENTION: 3-Phosphoadenosine-5-Phosphosulfate
; TITLE OF INVENTION: (PAPS) Synthetase Proteins and Methods for Treating
; TITLE OF INVENTION: Osteoarthritis Disorders
; FILE REFERENCE: 18810-81552
; CURRENT APPLICATION NUMBER: US/09/898,165B
; PRIOR FILING DATE: 2001-07-02
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1851
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-898-165B-10
Alignment Scores:
Pred. No.: 3.31e-32 Length: 1851
Score: 484.50 Matches: 101
Percent Similarity: 65.37% Conservative: 33
Best Local Similarity: 49.27% Mismatches: 56
Query Match: 26.53% Indels: 15
DB: 4 Gaps: 4
US-10-829-432-4 (1-343) x US-09-898-165B-10 (1-1851)
Qy 142 LysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeu 161
Db 43 AAATCCACCAATGTGGTCTACCAAGGCCCTCATGTGAGCAGGCAACAGAGACGACGATG 102

Qy 162 LeuGlyGln-----LysGlyCysValValTrpIleThrGlyLeuSerGlySer 177
Db 103 GTTGGACACCGGGGAGGATTCGAGGATGTACCGTGTGGCTAACAGGTCTCTCTGGTGT 162
Qy 178 GlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThr 197
Db 163 GGGAAAACCAACCAATAGCTTTGCTTTGGAAGATACCTTTGATCTTCACGCCATCCCATGT 222
Qy 198 TyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAla 217
Db 223 TACTCCCTGGATGGGACCAATGTCCGTCACTGSCCTTAATAAGAACCTGGGATTTCTCTGCC 282
Qy 218 GluAspArgAlaGluAsnIleArgValGlyGluValAlaLysLeuPheAlaAspAla 237
Db 283 GGGGACCCGAGAAGAGATATATCCGCGGATCCGCGAGGTGGCCAAAGCTCTTTGCCGACGCC 342
Qy 238 GlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAlaCysArg 257
Db 343 GGCCTGGTTGTCATCACCAGCTTTATCTCTCTTTTGCAAAAGGATCGTGAGAAATGCCGA 402
Qy 258 AlaLeu-----LeuProHisSerAsnPheIleGluValPheIleAspLeu 272
Db 403 AAATCCACGAATCAGCAGGACTCCG-----TTCCTTGAGATCTTTGTAGATCGG 453
Qy 273 ProLeuLysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGly 292
Db 454 CTTTAAATATCTGTGAAAGCCGAGACGTAAAGGACTCTACAAACGAGCCGACGACGGA 513
Qy 293 LysIleLysGlyPheThrGlyIleAspAspProTyrGluProPheIleAsnGlyGluIle 312
Db 514 GAGATTAAAGGTTTACAGGATCGATTCTGACTATGAGAAACCTTGAAATCCAGAGTGT 573
Qy 313 ValIleLysMetLysAspGluCysProSerProLysAlaMetAlaLysGlnValLeu 332
Db 574 GTGCTGAG-----ACCAACTTCTTCAGTAAGCGACTGTGTGCAACAGGTGTG 624
Qy 333 CysTyrLeuGluGlu 337
Db 625 GAACTTTTGCAGGAG 639
RESULT 10
US-09-898-165B-2
; Sequence 2, Application US/09898165B
; Patent No. 6818428
; GENERAL INFORMATION:
; APPLICANT: Daniel H. Cohn
; APPLICANT: Muhammad Faiyaz ul Haque
; APPLICANT: Lily M. King
; APPLICANT: Deborah Krakow
; TITLE OF INVENTION: 3-Phosphoadenosine-5-Phosphosulfate
; TITLE OF INVENTION: (PAPS) Synthetase Proteins and Methods for Treating
; TITLE OF INVENTION: Osteoarthritis Disorders
; FILE REFERENCE: 18810-81552
; CURRENT APPLICATION NUMBER: US/09/898,165B
; PRIOR FILING DATE: 2001-07-02
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-898-165B-2
Alignment Scores:
Pred. No.: 3.71e-32 Length: 2000
Score: 484.50 Matches: 101
Percent Similarity: 65.37% Conservative: 33
Best Local Similarity: 49.27% Mismatches: 56
Query Match: 26.53% Indels: 15
DB: 4 Gaps: 4

US-10-829-432-4 (1-343) x US-09-898-165B-2 (1-2000)

QY 142 LysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyClnSerAspArgGlnLysLeu 161
 Db 103 AATCCACCATGTGCTACACAGGCCCATCATGTGAGCAGGAACAGAGAGACAGTG 162
 QY 162 LeuGlyGln-----LysGlyCysValValTrpIleThrGlyLeuSerGlySer 177
 Db 163 GTTGGAAACACGAGGAGATTCCGAGGATGTACCGTGTGCTAACAGGTCTCTCTGTGT 222
 QY 178 GlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThr 197
 Db 223 GGGAAACACCAATAGCTTGTCTTGGAGAGTACCTTGTATCTCACGCCATCCCATGT 282
 QY 198 TyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAla 217
 Db 283 TACTCCTCGATGGGACCAATGTCCGTCAATGGCTTAATAGAACCTGGGATCTCTGCC 342
 QY 218 GluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAla 237
 Db 343 GGGGACCGAGAGAGAAATATCCGCGGATCGCGGAGGTGGCCAAAGCTCTTTGCCGAGCC 402
 QY 238 GlyValIleCysIleAlaSerLeuIleSerProTyrArgAspArgAspAlaCysArg 257
 Db 403 GGCCTGGTTGTCATCACAGCTTATCTCTCTCTTTTGCAGAGGATCGTGAGAAATGCCGA 462
 QY 258 AlaLeu-----LeuProHisSerAsnPheIleGluValPheIleAspLeu 272
 Db 463 AATATCCAGATCACAGGACTCCCG-----TTCTTGAGATCTTTGAGATGCG 513
 QY 273 ProLeuLysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGly 292
 Db 514 CCTTTAAATATCTGTGAAGCGAGACGTAAGAAAGACTCTACAAACGAGCCGAGCAGGA 573
 QY 293 LysIleLysGlyPheThrGlyLeuAspAspProTyrGluProProIleAsnGlyLysLeu 312
 Db 574 GAGATTAAAGGGTTTACAGGCATCGATTCGACTATGAGAAACCTGAAACTCCAGAGTGT 633
 QY 313 ValIleLysMetLysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeu 332
 Db 634 GTGCTGAAG-----ACCAACTGTCTTCAAGGCGACTGTGTCAACAGGTGGTG 684
 QY 333 CysTyrLeuGluGlu 337
 Db 685 GAATTTTTCAGGAG 699

RESULT 11

US-09-898-165B-9
 ; Sequence 9, Application US/09898165B
 ; Patent No. 6818428
 ; GENERAL INFORMATION:
 ; APPLICANT: Daniel H. Cohn
 ; APPLICANT: Muhammad Faiyaz ul Haque
 ; APPLICANT: Lily M. King
 ; APPLICANT: Deborah Krakow
 ; TITLE OF INVENTION: 3-Phosphoadenosine-5-Phosphosulfate
 ; TITLE OF INVENTION: (PAPS) Synthetase Proteins and Methods for Treating
 ; TITLE OF INVENTION: Osteoarthritis Disorders
 ; FILE REFERENCE: 18810-81552
 ; CURRENT APPLICATION NUMBER: US/09/898,165B
 ; CURRENT FILING DATE: 2001-07-02
 ; PRIOR APPLICATION NUMBER: 09/399,212
 ; PRIOR FILING DATE: 1999-09-17
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 1845
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-898-165B-9

Alignment Scores: 5.42e-32 Length: 1845

Pred. No.:

Score: 482.00 Matches: 98
 Percent Similarity: 66.67% Conservative: 30
 Best Local Similarity: 51.04% Mismatches: 52
 Query Match: 26.40% Indels: 12
 DB: 4 Gaps: 3

US-10-829-432-4 (1-343) x US-09-898-165B-9 (1-1845)

QY 133 GluLysProValMetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeu 152
 Db 13 AAGAAGCAAAAGACCGGAGAACACGAGAAATCCCAATGTAGTCTATCAGGCCCAACCAT 72
 QY 153 IleGlyGlnSerAspArgGlnLysLeuGlyGln-----LysGlyCysVal 168
 Db 73 GTGAGCAGGAATAAGAGAGGGCAAGTGTGGCAACAGGGGTGGTTCGAGGATGTACC 132
 QY 169 ValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAlaCysAlaLeuSerArg 188
 Db 133 GTGTGGCTAACAGAGTCTCTCTGTGTCTGAAAAACAACGATAAAGTTTTCCTCGAGGAG 192
 QY 189 GluLeuHisCysArgGlyHisLeuThrTyrValLeuAspGlyAspAsnLeuArgHisGly 208
 Db 193 TACCTGTCTCCCATCCCATCCCTGTACTCCCTGGATGGGACAAATGTCCGTATGCG 252
 QY 209 LeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIleArgArgValGly 228
 Db 253 CTTAACAGAAATCTCGGATCTCTCTCTGGGACAGAGAGAAATATCCGCCGATTGCT 312
 QY 229 GluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSerPro 248
 Db 313 GAGGTGGCTAAGCTCTTTGTCTGATCTGTCTGTCTGTCTGATCTTATTTCTTCTCCA 372
 QY 249 TyrArgAspArgAspAlaCysArgAlaLeu-----LeuProHisSer 263
 Db 373 TTCCGAAAGGATCGTGAGAAATCCCGCAAAATACATGATCATGACGAGGGCTGCCA 426
 QY 264 AsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProLys 283
 Db 427 ---TTCTTTGAAATATTGTAGATGCACCTCTAAATATTGTGAAAGCAGAGACGCTAAA 483
 QY 284 GlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyLysAspPro 303
 Db 484 GGCCTCTATAAAGGCGCAGAGCTGGGAGAGATTAAGGATTTACAGGATTTGATTTGAT 543
 QY 304 TyrGluProProIleAsnGlyGluIleValIleLys 315
 Db 544 TATGAGAAACCTGAAACTCCTGAGCGGTGTGCTTAAA 579

RESULT 12

US-09-153-310-41
 ; Sequence 41, Application US/09153310
 ; Patent No. 6326184
 ; GENERAL INFORMATION:
 ; APPLICANT: Gjermansen, Claes
 ; Hansen, Jorgen
 ; Johannesen, Pia Francke
 ; Pedersen, Mogens Bohl
 ; Sorensen, Steen Bech
 ; TITLE OF INVENTION: Method of producing a composite
 ; fermented beverage using genetically modified yeast
 ; strains
 ; NUMBER OF SEQUENCES: 42
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W.
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS


```
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/153,310
/ FILING DATE: 15-Sep-1998
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: <Unknown>
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bent, Stephen A.
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: <Unknown>
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-672-5300
/ TELEFAX: 202-672-5399
/ TELEX: <Unknown>
/ INFORMATION FOR SEQ ID NO: 41:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1160 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: Genomic DNA
/ SEQUENCE DESCRIPTION: SEQ ID NO: 41:
US-09-153-310-41
```

```
Alignment Scores:
Pred. No.: 1,1e-31 Length: 1160
Score: 475.00 Matches: 99
Percent Similarity: 66.85% Conservative: 24
Best Local Similarity: 53.80% Mismatches: 55
Query Match: 26.01% Indels: 6
DB: 3 Gaps: 3
```

US-10-829-432-4 (1-343) x US-09-153-310-41 (1-1160)

```
QY 143 SerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuLeu 162
   ::::::::::::::::::::
Db 381 GCTACTAATATCACTGGCATCCAAATCTTACC---TACGACGACGTAAGGAATTAAGA 437

QY 163 GlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeu 182
   ::::::::::::::::::::
Db 438 AAGCAAGACGGCTGTACCGCTTGGTGTGACCGGTCTAAGTGCCTCAGGAAAAAGTACAATA 497

QY 183 AlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAspGly 202
   ::::::::::::::::::::
Db 498 GCTGTGCACTGGAACAAATCTCTCTCAAAAAAACTTATCTGCTTATAGGTATAGTGT 557

QY 203 AspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGlu 222
   ::::::::::::::::::::
Db 558 GATAACATTCGTTTGGTTTGAATAAGGATTTGGGCTTCTCAGAAAAAGGACAGAAATGAA 617

QY 223 AsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIle 242
   ::::::::::::::::::::
Db 618 AACATTCGTAGATAGTAGCAAGTATCCAAAGCTATTGCGTGAATTCGTGTGCTGTATCCATC 677

QY 243 AlaSerLeuIleSerProTyrArgArgAspArgAlaCysArgAlaLeuLeuProHis 262
   ::::::::::::::::::::
Db 678 ACTTCATTTATTTCCCATCAGAGTCGATAGAGACAGAGCCGCGTGAATTTACATAAGNA 737

QY 263 Ser-----AsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArg 280
   ::::::::::::::::::::
Db 738 GCAGGGCTCAAGTTCAATGAAATTTTGTGTATGTTCCATTAGAAGTCGCTGAGCAAGA 797

QY 281 AspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIle 300
   ::::::::::::::::::::
Db 798 GACCCCTAAGGGTTGTATAGAAAGCCAGAGAGGGTGTGATTAAGAGGTTCAGTGTATT 857

QY 301 AspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAsp----- 318
   ::::::::::::::::::::
Db 858 TCAGCTCCTTACGAGCTCCAAAGGCCCGCAGAGTTGCTATTAAGAACTGACCCAAAGACT 917

QY 319 ----GluGluCys 321
```

```
Db 918 GTTGAAGAATGCT 929
RESULT 13
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENORU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1
```

```
Alignment Scores:
Pred. No.: 3,19e-26 Length: 640681
Score: 458.50 Matches: 99
Percent Similarity: 60.77% Conservative: 28
Best Local Similarity: 47.37% Mismatches: 75
Query Match: 25.11% Indels: 7
DB: 4 Gaps: 2
```

US-10-829-432-4 (1-343) x US-09-790-988-1 (1-640681)

```
QY 127 ProHisSerProValLysGluLysProValMetSerAsnIleGlyLys----- 142
   ::::::::::::::::::::
Db 457846 CCACATTTGGAACATA---CCAAAAAATTTAATGAAAAAGGTTTATAAGTAAATGAACAAT 457790

QY 143 -----SerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLys 160
   ::::::::::::::::::::
Db 457789 AATTTTCAAAACAAATATTTTGGCAAAAAACATTCAAATTACACGCTTAAAAACGTGAAAAA 457730

QY 161 LeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSer 180
   ::::::::::::::::::::
Db 457729 AAAAATGCTCATTAATCAATTTGTTACTATGGTTCTACTGGGCTCTCAGGTTTCAGAAAAATCA 457670

QY 181 ThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeu 200
   ::::::::::::::::::::
Db 457669 ACCATAGCAAAATTTTGTAGAAGAAATATTTGTTAAAAAATGGAATTAATAGCTATTATTA 457610

QY 201 AspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArg 220
   ::::::::::::::::::::
Db 457609 GACGAGACAAATATATAGGTCTGGTTTATGTTCTGATTTTAAGTTTTTAGTTTGGCAGATAGG 457550

QY 221 AlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIle 240
   ::::::::::::::::::::
Db 457549 AATGAAACAAATAGACGCAATTTGGAGAACTAGTAAAAATGATTTACATGCTGTTCTAATA 457490

QY 241 CysIleAlaSerLeuIleSerProTyrArgArgAspArgAlaCysArgAlaLeuLeu 260
   ::::::::::::::::::::
Db 457489 ATATTGGTATCAGTTATTTCCCTCTATAGAAATCAAAGGAAATGGTACGTCGTAATGTGA 457430

QY 261 ProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArg 280
   ::::::::::::::::::::
Db 457429 GGAIAAAAAAACTTTTATAGAAAGTTTTCATTGATACACCTATTGAATTTTGAAATCTCGA 457370

QY 281 AspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIle 300
   ::::::::::::::::::::
Db 457369 GATCCTAAAAAATATATAAAACAGCCGACAGCCCAAAATATCTGAATTTTACTGTTATA 457310

QY 301 AspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGluGlu 320
```



```
Db 457309 CAATGTACATACGAACTCCTATACACCTGATGTTCTTTAAAGGTTACAGATTCTTTA 457250
Qy 321 CysProSerProLysAlaMetAlaLys 329
Db 457249 AAAAATACTCAAAAAAATAATTAATAA 457223

RESULT 14
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 3,74e-20 Length: 4403765
Score: 402.50 Matches: 120
Percent Similarity: 44.38% Conservative: 34
Best Local Similarity: 34.58% Mismatches: 111
Query Match: 22.04% Indels: 83
DB: 3 Gaps: 11

US-10-829-432-4 (1-343) x US-09-103-840A-2 (1-4403765)
Qy 11 ProLeuValThrHisThrGlnGlnProSerProAlaProGlyProAlaSerGlnGly 30
Db 1439362 CCGCGCGGCGTACGTTGTCAAGCACCA-----CCGAACCGTCCGCGGAGGA 1439412
Qy 31 GlnArgGlnGlyAsnThrLeuLeuSerProThrProThrLeuAlaValileLeuValaen 50
Db 1439413 TAGCCGGGCTGGATTACCGGCTCGATGTCAACACCGCTGCATCGCG-----ACAAAGA 1439463
Qy 51 ProGlnArg-----AlaProProValLeu 58
Db 1439464 CCGCAACGGCGTTGAACCAAGCTGGCGGTGTTTCGTCGCACCGCCAGGTCGCGT 1439523
Qy 59 ProGlyLeuThrProSerAspAlaProLeuProAlaLeuValileHisGlyLeuThrPro 78
Db 1439524 TCGTGTTCAGAGTACACCGGACCGTAGCACCGGCTGCTTCATCTCATGACCCCG 1439583
Qy 79 ArgSerSerHisSerSerAlaGlyLeuAlaSerAspSerGlyArgArgGlyGlyGly 98
Db 1439584 ACACCAACGGA-----1439595
Qy 99 ArgGlyAlaArgThrHisCysHisArgGlyIleGlyArgTrrpValArgArgArgArg 118
Db 1439596 -----CGGTGG-----CGGCGG 1439607
Qy 119 AsnGlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSer 138
Db 1439608 GCATGTGTACCGGACGCTCGGCCCGC-AGCCT-----1439642
Qy 139 AsnIleGlyLeuSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyLysSerAspArg 158
Db -----1439690

Db 1439643 -----AGCCGAACACACCGTGGCGGACAGATCGCTCGTCACTGCGCAAGATCGG 1439690
Qy 159 GlnLysLeuGlyGlnLysGlyCysValValTrrpIleThrGlyLeuSerGlySerGly 178
Db 1439691 CCG-----CCGAGGGGCAAGACGGTGTGTTTACCGGACTGTCCGGCTCCGCG 1439738
Qy 179 LysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyr 198
Db 1439739 AAGTCGTGGTGGCCATGCTGTTGAGCGGAAGCTACTCGAAAGAGGATCTCCGCTTAC 1439798
Qy 199 ValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelysAlaGlu 218
Db 1439799 GTTCTGSAACGGGACCAACCTACGGCATGCGCTCAACGCGGACCTGGGCTTTTCCATGGCC 1439858
Qy 219 AspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGly 238
Db 1439859 GACCGCGGGAGAACCTGCGCGCTGTCGATGTGGGCACACTGCTGCCGATTTGTGGC 1439918
Qy 239 ValIleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAlaCysArgAla 258
Db 1439919 CACCTGGTGTGGTGGCGGATCAGCCCTTGTGAGCACCGTGGCTGGCTCGTAA 1439978
Qy 259 LeuLeuProHisSer-----AsnPhelleGluValPheIleAspLeuProLeu 274
Db 1439979 GTG-----CACGCTGATCGCGGAATCGACTTTTTCGAGGTGTTCTGACACCCCGGTG 1440032
Qy 275 LysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIle 294
Db 1440033 CAGGACTGTGAGAGCGGTGATCCCAAGGGTTGTACGCAAGCCGCGTGGGTGAGATC 1440092
Qy 295 LysGlyPheThrGlyIleAspAspProTyrGluProProIleAsnGlyGluIleValile 314
Db 1440093 ACGCACTTCACCGGGATCGACAGCCATATCAGCGGCCCCAAGAACCCAGACCTACGGCTT 1440152
Qy 315 -----LysMetLysAspGluCysProSerProLysAlaMetAlaLysGlnVal 331
Db 1440153 ACGCGGATCGCAGCATAGACGAGCAG-----GCCGAGGAGGTT 1440191
Qy 332 LeuCysTyrLeuGluGluAsn 338
Db 1440192 ATCGACCTGTTGGAGTCAATCG 1440212

RESULT 15
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 3,75e-20 Length: 4411529
Score: 402.50 Matches: 120
Percent Similarity: 44.38% Conservative: 34
Best Local Similarity: 34.58% Mismatches: 111
Query Match: 22.04% Indels: 83
DB: 3 Gaps: 11
```

US-10-829-432-4 (1-343) x US-09-103-840A-1 (1-4411529)

Search completed: September 9, 2005, 15:57:40
Job time : 2514 secs

```
QY 11 ProLeuValThrHisThrGlnProProSerProAlaProGlyProAlaSerGlnGly 30
Db 1439893 CCGGCGCGGACTACCTGTTCAAGCACACCA-----CCGAACCGTCCGCGCGAGGA 1439943
QY 31 GlnArgGlnGlyAsnThrLeuLeuSerProThrProThrLeuAlaValIleLeuValAsn 50
Db 1439944 TAGCCGGGCTGATTACCGGCTCGATGTCAACACCCTGCATCGCG-----ACAAGA 1439994
QY 51 ProGlnArg-----AlaProProValLeu 58
Db 1439995 CCGCAACGGCGTTGAAACTCAACGAACCTGGGCGGTGTTTCGCTGCGCACCCAGGTGCCGT 1440054
QY 59 ProGlyLeuThrProSerAspAlaProLeuProAlaLeuValIleHisGlyLeuThrPro 78
Db 1440055 TCCTGCTTACAGAGTACACCCGCAACGCTAGCACCGGCTCGTTCATCTCATTTGACCCCG 1440114
QY 79 ArgSerSerHisSerSerAlaGlyLeuAlaSerAspSerGlyArgGluGlyGly 98
Db 1440115 ACACCAACGGAA----- 1440126
QY 99 ArgGlyAlaArgThrHisCysHisArgGlyIleGlyArgTTPValArgArgArg 118
Db 1440127 -----CGGTGG-----CGGCGG 1440138
QY 119 AsnGlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSer 138
Db 1440139 GCATGGTGTACGCGACGCTCGGCCCGC-ACGCCT----- 1440173
QY 139 AsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArg 158
Db 1440174 -----AGCCCGAACCGGTGCGGCACACATCGCTCGTCACTGCGCAAGATCGG 1440221
QY 159 GlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGly 178
Db 1440222 CCG-----CCCAGGGGCAAGACGGTGTGTTTACCGGACTGTCCGGCTCCGCGC 1440269
QY 179 LysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyr 198
Db 1440270 AAGTCGTGCGTGGCCATGCTGTTGAGCGGAAGTACTCGAAAAGGGCATCTCCCGCTTAC 1440329
QY 199 ValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelysAlaGlu 218
Db 1440330 GTTCTGGACGGCGACAACTACGGCATGCGCTCAACGCGACCTGGGCTTTTCCATGGCC 1440389
QY 219 AspArgAlaGluAsnIleArgValGlyGluValAlaLysLeuPheAlaAspAlaGly 238
Db 1440390 GACCGCGCGAGAACCTGCGCGGCTGTGCGCATGTGGCCACACTGCTCGCCGATTGTGGC 1440449
QY 239 ValIleCysIleAlaSerLeuIleSerProTyrArgAspArgAspAlaCysArgAla 258
Db 1440450 CACCTGGTGTGTCGCCGATACGCCCTTGTGTCAGCACCGTCCCTGGCTGTGCTGTA 1440509
QY 259 LeuLeuProHisSer-----AsnPheIleGluValPheIleAspLeuProLeu 274
Db 1440510 GTG-----CACGCTGATCGGGGAATCGACTTTTTCGAGGTGTCTGTGACACCCCGCTG 1440563
QY 275 LysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIle 294
Db 1440564 CAGGACTGTGAGAGGCGTATCCCAAAGGGTTGTACGCCAAAGCGCGTGGGTGAGATC 1440623
QY 295 LysGlyPheThrGlyIleAspAspProTyrGluProProIleAsnGlyGluIleValIle 314
Db 1440624 ACGCACTTACCGGGATCAGACGCCATATACGGCCCAAGAACCCAGACCTACGGCTT 1440683
QY 315 -----LysMetLysAspGluGluCysProSerProLysAlaMetAlaLysGlnVal 331
Db 1440684 ACGCCGGATCGCAGCATACGAGCAG-----GCCGAGGAGGTT 1440722
QY 332 LeuCysTyrLeuGluAsn 338
Db 1440723 ATCGACCTGTGGAGTCATCG 1440743
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2005, 12:40:30 ; Search time 47 Seconds
(without alignments)
702.178 Million cell updates/sec

Title: US-10-829-432-4
Perfect score: 1826
Sequence: 1 RPFHFINQTEPLVTHTQPP.....PRMAKQVLCYLENGYLQA 343

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	833	45.6	312	2 T08076	adenyl-yl-sulfate k
2	816	44.7	276	1 S47640	adenyl-yl-sulfate k
3	777.5	42.6	293	2 T08100	adenyl-yl-sulfate k
4	571	31.3	200	2 E96912	adenyl-yl-sulfate ki
5	555	30.4	199	2 A69839	adenyl-yl-sulfate ki
6	534.5	29.3	202	2 A83836	adenyl-yl-sulfate ki
7	511.5	28.0	633	2 H95932	probable adenyl-yl-
8	510	27.9	213	2 AF0408	adenyl-yl-sulfate k
9	508.5	27.8	635	2 A87433	hypothetical prote
10	505	27.7	652	1 T24918	3'-phosphoadenosin
11	503.5	27.6	202	2 T50101	adenyl-yl-sulfate ki
12	502.5	27.5	610	1 JC4383	adenyl-yl-sulfate k
13	498.5	27.3	660	2 G82672	ATP sulfurylase 1
14	495	27.1	201	2 AH0856	adenosine 5'-phosp
15	495	27.1	624	1 JW0087	adenyl-yl-sulfate k
16	494.5	27.1	215	2 F82062	adenyl-yl-sulfate k
17	492	26.9	644	2 AD3471	adenyl-yl-sulfate k
18	490.5	26.9	641	1 Z22RNO	adenyl-yl-sulfate k
19	490.5	26.9	641	2 E95320	adenyl-yl-sulfate k
20	481	26.3	201	1 B65056	adenyl-yl-sulfate k
21	481	26.3	201	2 D91079	adenosine 5' -phosp
22	481	26.3	201	2 E85924	adenosine 5' -phosp
23	481	26.3	197	2 A84073	adenyl-yl-sulfate ki
24	480	26.3	198	2 C69877	adenyl-yl-sulfate k
25	478.5	26.2	620	1 I39755	adenyl-yl-sulfate k
26	474	26.0	202	1 S17244	adenyl-yl-sulfate k
27	457	25.0	206	2 H84978	adenyl-yl-sulfate k
28	448.5	24.6	196	2 H83472	adenosine 5' -phosp
29	402	22.0	614	1 B70772	probable adenyl-yl-

30	398	21.8	546	1 C70393	probable adenyl-yl-
31	387	21.2	177	1 S74917	adenyl-yl-sulfate k
32	369	20.2	214	1 S18729	adenyl-yl-sulfate k
33	351	19.2	192	2 B75594	adenyl-yl-sulfate ki
34	325	17.8	186	2 G72590	probable adenyl-yl-
35	322.5	17.7	574	1 S55034	sulfate adenyl-yltr
36	314.5	17.2	174	2 F75097	adenyl-yl-sulfate 3-
37	307	16.8	573	1 A53651	sulfate adenyl-yltr
38	305	16.7	155	2 H69285	adenyl-yl-sulfate 3-
39	302	16.5	633	2 D83091	ATP sulfurylase GT
40	179	9.8	170	2 G81286	probable adenyl-yl-
41	111.5	6.1	396	2 T36678	probable septum si
42	108	5.9	1016	2 D86308	translation initia
43	106	5.8	478	2 T10030	hypothetical prote
44	106	5.8	586	2 H86914	conserved hypothet
45	106	5.8	723	2 B38749	3-phosphatidylnos

ALIGNMENTS

RESULT 1

T08076
adenyl-yl-sulfate kinase (EC 2.7.1.25) precursor - Madagascar periwinkle
N:Alternate names: adenosine-5'-phosphosulfate-kinase
C:Species: Catharanthus roseus (Madagascar periwinkle)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 09-Jul-2004
C:Accession: T08076
R:Arz, H.E.; Gisselmann, G.; Schiffmann, S.; Schwenn, J.D.
Biochim. Biophys. Acta 1218, 447-452, 1994
A:Title: A cDNA for adenyl-yl sulphate (APS)-kinase from Arabidopsis thaliana.
A:Reference number: S47640; MUID:94325358; PMID:8049272
A:Accession: T08076
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-312 <AR>
A:Cross-references: UNIPROT:O49204; EMBL:AF044285; NID:g28322299; PIDN:AAC31145.1; PID:g2
C:Genetics:
A:Gene: akn
C:Function:
A:Description: phosphorylates 3'-OH group of adenosine-5'-phosphosulfate
C:Superfamily: adenyl-yl-sulfate kinase; adenyl-yl-sulfate kinase homology
C:Keywords: ATP; P-loop; phosphotransferase; purine nucleotide binding
F:135-298/Domain: adenyl-yl-sulfate kinase homology <ASK>

Query Match 45.6%; Score 833; DB 2; Length 312;
Best Local Similarity 72.9%; Pred. No. 2.2e-55;
Matches 156; Conservative 24; Mismatches 34; Indels 0; Gaps 0;

Qy	130	PVKEKPVMSNIGKSTNIIWHNCLIGQSDRQKLLGQKGVWITGLSGSGKSTLACALSRG	189
Db	99	PGKILQTTTGVNSTNIIWHKCAVEKSERQEPLOQRGCVIWTGLSGSGKSTLACALSRG	158
Qy	190	LHCRGHLYVLDDGNLRHGLNRDLSFKAEDRAENIRRVGEVAKLPADAGVTCIASLISPY	249
Db	159	LHAKGKLYI LDGDNVVRHGLNSDLSFKAEDRAENIRRVGEVAKLPADAGVTCIASLISPY	218
Qy	250	RRDRDACPALLPHSNFIEVFIDLPLKICEARDPKLYKLARTGKTKGTGDDDPYEPPIP	309
Db	219	RKPPDACRSLPEGDFIEVFMDVPLKVCEARDPKLYKLARAGIKGTGIDDDPYEPPLK	278
Qy	310	GEIVKMKDECPSPKAMAKOVLCVLENGYLQA	343
Db	279	SEIVLHQKLGMCDSPCDLADIVISYLENGYLKA	312

RESULT 2

S47640
adenyl-yl-sulfate kinase (EC 2.7.1.25) precursor - Arabidopsis thaliana
N:Alternate names: APS kinase; protein F26C24.11
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S47640; T02601; A84521; S38587

A;Status: preliminary; translated from GB/EMBL/DDBB
A;Molecule type: mRNA
A;Residues: 1-293 <AR2>
A;Cross-references: EMBL:AF043351; PIDN:AAC39520.1
C;Genetics:
C;Gene: akn2
A;Map position: 4
A;Introns: 78/1; 177/1; 216/3; 245/1; 260/3
A;Note: T5J17.110
C;Function:

Query Match	42.6%;	Score	777.5;	DB	2;	Length	293;
Best Local Similarity	71.8%;	Pred. No.	3.1e-51;				
Matches	145;	Conservative	26;	Mismatches	30;	Indels	1;
						Gaps	1;

[illegible]

Db	211	DGDFVEFMDVPLHVCESDRDPKGLYKLRAGKIKGFTGIDDDPYEAPUNCCEVVLKHTGDDE	2170
Qy	321	CPSPKAMAKQVLCYLBENGYLQ	342
Db	271	SCSPROMAENIISYLNKGYLE	292

Db 271 SCSPROMAENIISYLNKGYLE 292

adenylylsulfate kinase [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
E96912
NCBI

C/Date: 14-Sep-2001 #session_revision 14-Sep-2001 #teac_charge 09-Mar-2004
C/Accession: E96912
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson,
D.; Dalv, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4836, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clostridium*
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: E96912

A;Status: preliminary
A;Molecule type: DNA
A;Molecular weight: 1-200 <KUR>
A;Residues: 1-200 <KUR>
A;Cross-references: UNIPROT:O97MT8; GB:AAK78088.1; PID:c15023928

C;Genetics: A;Gene: CAC0103
C;Superfamily: adenylylsulfate kinase; adenylylsulfate kinase homology

Query Match 31.3%; Score 571; DB 2; Length 200;
Best Local Similarity 54.9%; Pred. No. 6.8e-36;
Matches 113; Conservative 35; Mismatches 48; Indels 10; Gaps 3;
C:superfamily: adenylisulfate kinase; adenylisulfate kinase homolog

Query Match 31.3%; Score 571; DB 2; Length 200;
Best Local Similarity 54.9%; Pred. No. 6.8e-36;
Matches 113; Conservative 35; Mismatches 48; Indels 10; Gaps 3;

1


```
Db 549 RYAEVARLMDAGLVLVSVFSPFDRQLARERFAANFEVFDVPLVAEARDVKGL 608
QY 286 VKLARTGKIGFTGDDPYEPPIGEIVIKMDKDECPSPKMAKQVLCYL 335
Db 609 YAKARAGLITDTGIDSPYPPHPHDLUR---ADQGTPEQLASQVLSFL 655

RESULT 14
AH0856
adenosine 5-phosphosulfate kinase [imported] - Salmonella enterica serov
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AH0856
R.Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AH0856
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-201 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD06039.1; PID:gl6504006; GSPDB:GN00176
A:Gene: cysC
C:Genetics:
C:Superfamily: adenylylsulfate kinase; adenylylsulfate kinase homology

Query Match 27.1%; Score 495; DB 2; Length 201;
Best Local Similarity 54.0%; Pred. No. 3.6e-30;
Matches 95; Conservative 29; Mismatches 46; Indels 6; Gaps 1;

QY 145 NTLHNCILIGSDRQKLLGQGVVITGLSGSGKSTLACALSRELHCRGHLYVLDGDN 204
Db 7 NVVHSHPTVAAREQLGHGVRVLTFTGLSGSGKSTVAGALEALHQGVSTYLLDGDN 66

QY 205 LRHGLNRDLSPKAEDRAENIRRVGEVAKLPADAGVICTASLISPYRRDRACRALLPHSN 264
Db 67 VHHGLCRDLGFSADQRQENIRRVGEVASLMDAGLIVLTAFISPHRAERQLKRVGHDR 126

QY 265 FIEVDLPLKICEARDPKGLKARTGKIGFTGDDPYEPPIGEIV 314
Db 127 FIEIVNTPLAICEQRDPKGLYKARAGELRNFTGIDAIYEAPDSPQVHLNGEQLV 182

RESULT 15
JW0087
adenylyl-sulfate kinase (EC 2.7.1.25) - human
N:Alternate names: adenosine 5'-phosphosulfate kinase; PAPS
N:Contains: adenylylsulfate kinase (EC 2.7.1.25); sulfate adenylyltransferase (EC 2.7.7.
C:Species: Homo sapiens (man)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: JW0087
R.Yanagisawa, K.; Sakakibara, Y.; Suiko, M.; Takami, Y.; Nakajima, H.; Tak
Biosci. Biotechnol. Biochem. 62, 1037-1040, 1998
A:Title: cDNA cloning, expression, and characterization of the human bifunctional ATP su
A:Reference number: JW0087; MUID:98312048; PMID:9648242
A:Accession: JW0087
A:Molecule type: mRNA
A:Residues: 1-624 <VAN>
A:Cross-references: UNIPROT:O43252; GB:AF033026; NID:g3378100; PIDN:AAC28429.1; PID:g337
A:Experimental source: Brain
C:Function: <ASKF>
A:Description: as adenylylsulfate kinase catalyzes the phosphorylation of adenylylsulfat
C:Function: <SATF>
A:Description: as sulfate adenylyltransferase catalyzes the reaction of sulfate and ATP
C:Superfamily: animal 3'-phosphoadenosine-5'-phosphosulfate synthetase; adenylylsulfate
C:Keywords: multifunctional enzyme; nucleotide binding; nucleotidyltransferase; P-loop;
F:52-215/Domain: adenylylsulfate kinase homology <ASK>
F:59-66/Region: nucleotide-binding motif A (P-loop)
```

```
F:226-620/Domain: sulfate adenylyltransferase homology <SAT>
F:133/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 27.1%; Score 495; DB 1; Length 624;
Best Local Similarity 46.1%; Pred. No. 1.4e-29;
Matches 106; Conservative 38; Mismatches 60; Indels 26; Gaps 6;

QY 125 EAPHSPVKEKPMNSN-----IGKSTNLIWHNCLIGSDRQKLLGQ-----KGVVVIITGLS 175
Db 2 EIPGSLCKVKVLSNNAQWGMORATNTVYQAHVSRNKRQGVVTRGGFRGCTVWLTGLS 61

QY 176 GSGKSTLACALSRELHCRGHLYVLDGDNLRHGLNRDLSPKAEDRAENIRRVGEVAKLPA 235
Db 62 GAGKTTVSMALVEYLCHGIPCYTLGDGNIHQGLNKLGFSPDEENVRRIAEVAKLPA 121

QY 236 DAGVICTASLISPYRRDRDADR-----ALLPHSNFIEVFDLPLKICEARDPKGLYKLR 290
Db 122 DAGLVCTISFTSPYTDQRNNARQTHEGASLP---FFEVFDAPLHVCEQRDVKGLYKKAR 178

QY 291 TGKIKGFTGIDDPVEPPINGEIVIKMD---EECPSPKAMAKQVLCYLEE 337
Db 179 AGEIKGFTGIDSEYEKPEAPELVLTDSVDYDC-----VQQVVELLOE 222
```

Search completed: September 9, 2005, 13:04:01
Job time : 48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 9, 2005, 13:04:07 ; Search time 3067 Seconds
(without alignments)
4256.943 Million cell updates/sec

Title: US-10-829-432-4

Perfect score: 1826

Sequence: 1 RPFHINOTEPLVTHQPPP.....PRMAKQVLCYLENGYLQA 343

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool_p/US10829432/runat_08092005_132223_25088/app_query.fasta_1.519
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -TRANS=human40.cdi -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10829432 @CNG_1_1_5180 @runat_08092005_132223_25088 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1246	68.2	720	7	COS28976
2	1080.5	59.2	645	7	COS32501
3	944	51.7	915	9	CL981208
4	936	51.3	675	2	BE357876
5	926	50.7	691	2	BE355111
6	881.5	48.3	1075	7	CK205911
7	874	47.9	614	6	CA204452
8	865	47.4	617	6	CD863007
9	863.5	47.3	860	7	CR285854

10	837.5	45.9	673	6	CA182186
11	816.5	44.7	867	7	CO095086
12	816	44.7	1066	7	CNS0A87V
13	810.5	44.4	991	3	AY110087
14	801.5	43.9	659	6	CA900730
15	798	43.7	970	3	AY109629
16	797	43.6	555	4	BI427055
17	797	43.6	809	6	CA766480
18	793	43.4	696	6	CA085472
19	792.5	43.4	837	7	CO200697
20	789	43.2	807	7	CNI34308
21	775	42.4	797	7	CNI29765
22	774.5	42.4	1206	3	CNS0A2AD
23	773	42.3	754	6	CA921172
24	772	42.3	1090	5	EX838386
25	760.5	41.6	648	6	CD817873
26	757.5	41.5	1163	3	CNS091ZG
27	751	41.1	536	2	AW560397
28	748.5	41.0	1146	3	CNS0A2PA
29	748	41.0	525	6	CA502007
30	747	40.9	595	5	BU808718
31	744	40.7	872	7	CV281393
32	743	40.7	673	6	CD924390
33	738	40.4	496	6	CA497171
34	735	40.3	902	7	CV245034
35	733.5	40.2	748	4	BG126657
36	728	39.9	564	2	AW056154
37	728	39.9	957	7	CV260804
38	725	39.7	860	7	CF515151
39	723	39.6	700	1	AJ502331
40	717	39.3	803	7	CO476247
41	717	39.3	814	6	CA759382
42	714	39.1	1158	3	CNS09ZLM
43	712	39.0	772	2	BF631366
44	711	38.9	846	7	CF515003
45	710	38.9	617	5	BQ518805

ALIGNMENTS

RESULT 1

COS28976

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COS28976 720 bp mRNA linear EST 15-JUL-2004
3530.1.191.1.D07.Y.1.3530 - Full length cDNA library created by
Invitrogen from multiple tissues Zea mays cDNA, mRNA sequence.

COS28976.1 GI:50333850

EST.

Zea mays

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 720)

Walbot V.

Maize ESTs from various cDNA libraries sequenced at Stanford

University

Unpublished (1999)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 3530.1.191.1 row: D column: 07.

Location/Qualifiers

1..720

/organism="Zea mays"

/mol_type="mRNA"

/cultivar="B73"

/db_xref="taxon:4577"

/tissue_type="multiple"

/dev_stages="varies by tissue"
 /lab_host="DH10B"
 /clone_lib="3530 - Full length cDNA library created by
 Invitrogen from multiple tissues"
 /note="Organ: silks, husks, ears, pollen, shoot tips,
 leaf, root tips, whole seed, embryo; Vector: pCMV-SF0RT
 6.1; Site 1: EcoRV; Site 2: NotI; Maize Gene Discovery
 Project contracted with Invitrogen to produce a
 normalized, full length library in a pSport vector. This
 is a Gateway compatible vector, permitting clone movement
 to new vector backbones for expression in diverse host
 cells using recombination rather than restriction enzymes.
 Details of the vector and sequencing primers are available
 at ZmDB in the EST library description tables. poly(A)+
 mRNA was prepared by Invitrogen, and equimolar amounts of
 RNA from each of the 12 tissue samples were mixed together
 for selection of mRNA with a 5' cap. After synthesis of
 cDNA, a normalization step was conducted against the
 mixture of RNA sources. This step effected a 20X to 80X
 reduction in common transcript types. Tissues prepared: 1.
 just emerging silks; 2. inner husks from ears of sample
 #1; 3. 20 dap aleurone; 4. immature tassels, stages from
 1-2 mm to 1-2 cm; 5. 2 mm to 2 cm ears; 6. pollen; 7. 1 cm
 vegetative shoot tips from 15 day old seedlings; all
 leaves with an expanded or partially expanded sheath
 were removed; 8. mature leaf tissue; 9. 0.5 cm long root
 tips from 15 day old seedlings; 10. 10 dap whole seed; 11.
 12 dap endosperm and embryo; 12. 17 dap endosperm and
 embryo. All of the sequenced clones in project 3530 will
 be archived at the University of Arizona along with the
 Unigene clones from the Maize Gene Discovery EST
 sequencing projects. Clones can be ordered through the
 ZmDB web site or directly from the University of Arizona
 (<http://www.genome.arizona.edu/orders/>). High density
 filters containing over 18,000 clones can also be ordered
 from the University of Arizona."

ORIGIN

Alignment Scores:
 Pred. No.: 1,15e-87 Length: 720
 Score: 1246.00 Matches: 238
 Percent Similarity: 99.58% Conservative: 1
 Best Local Similarity: 99.17% Mismatches: 0
 Query Match: 68.24% Indels: 1
 DB: 7 Gaps: 0

US-10-829-432-4 (1-343) x COS28976 (1-720)

QY	28	SerGlnGlyGlnArgGlnGlyAsnThrLeuLeuSerProThrProThrLeuAlaValle	47
Db	2	AGCCAGGGCCAAACGGCAAGCAACACCTCTCTCAGCCGACGCCGCTGCGGTCATC	61
QY	48	LeuValAsnProGlnArgAlaProValLeuProGlyLeuThrProSerAspAlaPro	67
Db	62	CTCGTAATCCACAGCGCGCGCTCCCGTCCTCCAGCGCTCACCCCTAGCGATCGCCA	121
QY	68	LeuProAlaLeuValIleHisGlyLeuThrProArgSerSerHisSerAlaGlyLeu	87
Db	122	CTCCCGGCGCTGATCCATGGCTCACTCCCGCTTCTCACACTTCCGCGGCTCTC	181
QY	88	AlaSerAspSerGlyArgArgGluGlyGluGlyArgGlyAlaArgThrHisCysHisArg	107
Db	182	GCCAGCGATAGTGGCGCGCGCGAGGGAGGGCGCGGTGCGGTACGCACTCCACCGC	241
QY	108	GlyIleGlyArgTrpValArgArgArgArgAsnGlyAlaAlaPro-GlyGluAlaPr	127
Db	242	GGCATTTGGCGTGGTGGCGCGCGCGCGGATGGAGAGCCACGGGGAGCCCC	301
QY	127	HisSerProVallysGluLysProValMetSerAsnIleGlyLysSerThrAsnIleLe	147
Db	302	GCACAGCCAGTGAAGGAGAACCTGTATGTCAACATTTGGAAATCGACTAATATT	361
QY	147	utrHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuLeuGlyGlnLysGlyCy	167

Db	362	ATGCCACAATTCCTTGATGGACAATCTGATAGACAAAATTCGTGGCACAAGGCTG	421
QY	167	sValValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAlaCysAlaLeuSe	187
Db	422	TGTGTAATGGATAACAGGACTCAGTGGTTACGGGAAAGTACTCTTGCATGTGACTGAG	481
QY	187	rArgGluLeuHisCysArgGlyHisLeuThrTyValLeuAspGlyAspAsnLeuArgHi	207
Db	482	TCGTGAGTTGTCATTGCAGAGGCCACCTCACGTATGTACTTGTATGGTGACCACTCAGACA	541
QY	207	sGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIleArgArgVa	227
Db	542	TGGCCTTAATAGAGATTAAAGCTTTAAGGCGCAGACCGTCAGAAAATATACGAAGAGT	601
QY	227	lGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSe	247
Db	602	TGGTGAATGGCAAGCTTTTCGCTGATGCTGGTGTCTATGTCATTGCTAGCTTGATATC	661
QY	247	rProTyArgArgAspArgAspAlaCysArgAlaLeuLeuProHisSerAsnPheIle	266
Db	662	TCCATACAGGAGAGATCGTGATGATGCGCGTCTACTTCCACATTTACTTTATT	719

RESULT 2
 COS32501
 LOCUS
 DEFINITION 3530.1.213.1.H05.Y.1.3530 - Full length cDNA library created by
 Invitrogen from multiple tissues Zea mays cDNA, mRNA sequence.
 ACCESSION COS32501
 VERSION COS32501.1 GI:50337375
 KEYWORDS EST.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (Bases 1 to 645)
 AUTHORS Walbot,V.
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
 UNIVERSITY
 JOURNAL Unpublished (1999)
 COMMENT Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 3530.1.213.1 row: H column: 05.
 Location/Qualifiers
 1..645
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /tissue_type="multiple"
 /dev_stage="varies by tissue"
 /lab_host="DH10B"
 /clone_lib="3530 - Full length cDNA library created by
 Invitrogen from multiple tissues"
 /note="Organ: silks, husks, ears, pollen, shoot tips,
 leaf, root tips, whole seed, embryo; Vector: pCMV-SF0RT
 6.1; Site 1: EcoRV; Site 2: NotI; Maize Gene Discovery
 Project contracted with Invitrogen to produce a
 normalized, full length library in a pSport vector. This
 is a Gateway compatible vector, permitting clone movement
 to new vector backbones for expression in diverse host
 cells using recombination rather than restriction enzymes.
 Details of the vector and sequencing primers are available
 at ZmDB in the EST library description tables. poly(A)+
 mRNA was prepared by Invitrogen, and equimolar amounts of
 RNA from each of the 12 tissue samples were mixed together
 for selection of mRNA with a 5' cap. After synthesis of

cDNA, a normalization step was conducted against the mixture of RNA sources. This step effected a 20X to 80X reduction in common transcript types. Tissues prepared: 1. just emerging silks; 2. inner husks from ears of sample #1; 3. 20 dap aleurone; 4. immature tassels, stages from 1-2 mm to 1-2 cm; 5. 2 mm to 2 cm ears; 6. pollen; 7. 1 cm vegetative shoot tips from 15 day old seedlings; all leaves with an expanded or partially expanded sheath were removed; 8. mature leaf tissue; 9. 0.5 cm long root tips from 15 day old seedlings; 10. 10 dap whole seed; 11. 12 dap endosperm and embryo; 12. 17 dap endosperm and embryo. All of the sequenced clones in project 3530 will be archived at the University of Arizona along with the Unigene clones from the Maize Gene Discovery EST sequencing projects. Clones can be ordered through the ZmDB web site or directly from the University of Arizona (<http://www.genome.arizona.edu/orders/>). High density filters containing over 18,000 clones can also be ordered from the University of Arizona."

ORIGIN

Alignment Scores:

Pred. No.: 9,43e-75 Length: 645
Score: 1080.50 Matches: 213
Percent Similarity: 81.68% Conservative: 1
Best Local Similarity: 81.30% Mismatches: 0
Query Match: 59.17% Indels: 48
DB: 7 Gaps: 1

US-10-829-432-4 (1-343) x COS32501 (1-645)

QY 22 ProAlaProGlyProAlaSerGlnGlyGlnArgGlnGlnValHisThrLeuLeuSerProThr 41
DB 2 CCAGCGCGCGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 61
QY 42 ProThrLeuAlaValLeuValAsnProGlnArgAlaProProValLeuProGlyLeu 61
DB 62 CCGAGCGCTCGCGCTCATCTCTGTAATCCACAGCGCGCGCTCCCGTCTCTCCAGCGCTC 121
QY 62 ThrProSerAspAlaProLeuProAlaLeuValHisGlyLeuThrProArgSerSer 81
DB 122 ACCCTTAGCGATGCGCCATCTCCCGCGCTCGTGATCCATGGCTCACTCCCGCTTCCTCA 181
QY 82 HisSerSerAlaGlyLeuAlaSerAspSerGlyArgGlyGlyGlyValGlyValGlyAla 101
DB 182 CACTTTTCGCGGGTCTCGCCAGCGATAGTGGCGCGCGCGAGCGCGCGCGCGTGG 241
QY 102 ArgThrHisCysHisArgGlyIleGlyArgGlyValArgArgArgArgAsnGlyAla 121
DB 242 CGTACGCACTGCGACCGCGCATTTGGCGGTGGTGGTGGCGCGCGCGCGGAATGGAGCA 301
QY 122 AlaPro-GlyGluAlaProHisSerProValLysGluLysProValMetSerAsnIleG1 141
DB 302 GCGCCACGGGATGCGCGCCAGCAGCCCA----- 329
QY 141 yLysSerThrAsnIleLeuThrHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLe 161
DB 329 ----- 329
QY 161 uLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerTh 181
DB 330 -----GGGAAAAGTAC 340
QY 181 rleuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrrValLeuAs 201
DB 341 TCTTCATGTGCACTGAGTCGAGTCGAGTCGATTCGATTCGAGAGCCACCTTCACGTATGAT 400
QY 201 pGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAl 221
DB 401 TGGTACAAACCTTCACATATGGCTTAATAGATTTAAGCTTTAAGGCAGAGACCGTGC 460
QY 221 aGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCy 241

DB 461 AGAAAAATATACGAAGAGTTGGTGAAGTGGCAAAGCTTTTCGCTGATGCTGTGTCATATG 520
QY 241 sileAlaSerLeuIleSerProTyraArgAspArgAspAlaCysArgAlaLeuLeuPr 261
DB 521 CATTCCTAGCTTGATATCTCCATACAGAGAGATCGTGATGATGCCGTCTACTTCC 580
QY 261 oHisSerAsnPhelIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAs 281
DB 581 ACAITCTAACTTATTGAAGTATTATTGATTGCCCCCTAAAAATTTGTGAAGCTCGTGA 640
QY 281 pPro 282
DB 641 TCCT 644
RESULT 3
CL981208
LOCUS CL981208
DEFINITION OsIFCC035825 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.
ACCESSION CL981208
VERSION CL981208.1 GI:52416908
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 915)
AUTHORS Ma,L., Wang,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M., Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L., Wong,G.K.S., Deng,X.W. and Wang,J.
TITLE An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis
JOURNAL Unpublished (2004)
COMMENT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-8048676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
source Location/Qualifiers
1..915
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Express Library"
/note="Oryza sativa exon trapped genomic sequences"
ORIGIN
Alignment Scores:
Pred. No.: 7.73e-64 Length: 915
Score: 944.00 Matches: 194
Percent Similarity: 74.01% Conservative: 31
Best Local Similarity: 63.82% Mismatches: 59
Query Match: 51.70% Indels: 20
DB: 9 Gaps: 5
US-10-829-432-4 (1-343) x CL981208 (1-915)
QY 58 LeuProGlyLeuThrProSerAspAlaProLeuProAlaLeuValIleHis----- 74
DB 5 TTGCAGGAGATCTCCCATCCCGTCTCCGCGCGCGCGCGCGCGCGCGCGGTG 64
QY 75 -----GlyLeuThrPro-ArgSerSerHisSerSerAlaGlyLeuAlaSer----- 89
DB 65 CTCGCGATGGCGCTCGCGGAGGGGTACTAGAGCCACCGTGGGAGGAGTAGTGGCG 124
QY 90 -AspSerGlyArgArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 109
DB 125 GCGTGAGGCTCGTCGAGCGCGAGGCAATGAGCGTCTGTCAGCGCGCGCGCGCGCA- 183

QY	109	leGlyargTrpValArgArgArgArgAsnGlyAlaAa---	ProGlyyGluAlaProH	128
DB	184	--AGCAGCTGGGACAGGAAGCTGC	CGGTGGAGGCGCGCTCGCGCGGCGCGCGCG	241
QY	128	isSer-----	ProValLysGluLysProValMetSerAsnI	140
DB	242	TCGACGGGAAGGACGACGACGATGTTG	CGCAGTGTAAACAAGCTCATGACCTCAACTG	301
QY	140	leGlyLysSerThrAsnIleLeuTPhisAsnCysLeuIleGlyGlnSerAspArgGlnL	160	
DB	302	TCGGTAATAACAAACATCCGATGGCATGCTCC	CAGTAACACCAAGCTTGCACAGACAA	361
QY	160	ysLeuLeuGlyGlnLysGlyCysValValT	PhIleThrGlyLeuSerGlySerGlyLys	180
DB	362	AGTTGCTGACCAAGAGGCTGTGTGTTCGATCACT	TGGCTTAAGTGTTCAGGAANA	421
QY	180	erThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValL	200	
DB	422	GCACCCCTTGTCATGTCCACTGAGCCGTGAGCTGCACTCA	GAGGGCATCTGACCTATGTT	481
QY	200	euAspGlyAspAsnLeuHrgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspA	220	
DB	482	TTGACGGCGACAAATCTCCGGCACGGCCTTGAA	CAAAAGATCTCAGCTTCAAAAGCCACAGGATC	541
QY	220	rgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValI	240	
DB	542	GTGCGGAAATATACGAGAGTTGGAGAGTGGCAAGCTGT	TTGCAAGATGCTGGATATGA	601
QY	240	leCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeuL	260	
DB	602	CTGTGCTACTAGTTTGTATACCCATATAGAGTGATCGA	AGCGCTCGCGCAAAATAC	661
QY	260	euProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaA	280	
DB	662	TACCCAATTCCTCGTTTCATTAGGTGTTCCTGGAATG	TCCCATTTGAAGTATGTGAAGAA	721
QY	280	rgAspProlysglyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrClyT	300	
DB	722	GGGATCAAAAGCGCTGTACAAGCTTGCTCGTGGCGCAAAATCA	AAAGGCTTTTACGGGAA	781
QY	300	leAspAspProTyrGluProProfileAsnGlyGluIleValIleLysMetLysAspGluC	320	
DB	782	TAGATGATCCTTATGAAACACCTTCAGATTCGAGATTGTGATACAGTGC	CAAAAGTTGGGG	841
QY	320	lyCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGlyT	340	
DB	842	ACTGCCCTTCACCTAAATCAATGGCTGATCAAGTAGTGTGCATATCT	TTTGAAGCCCAATGGAT	901
QY	340	yrLeuGln	342	
DB	902	TCCTTCAG	909	

RESULT 4	BE357876	DG1_22_A06_g1_A002	Dark Grown 1 (DG1)	Sorghum bicolor	CDNA, mRNA	675 bp	mRNA	linear	EST 20-JUL-2000
LOCUS	BE357876								
DEFINITION	BE357876								
ACCESSION	BE357876								
VERSION	BE357876.1	GI:9299433							
KEYWORDS	EST.								
SOURCE	Sorghum bicolor	(sorghum)							
ORGANISM	Sorghum bicolor								
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.								
REFERENCE	1 (bases 1 to 675)								
AUTHORS	Cordonnier-Pratt,M.-M., Gingie,A., Marsala,C., Sudman,M. and Pratt,L.H.								
TITLE	An EST database from Sorghum: dark-grown seedlings								
JOURNAL	Unpublished (2000)								
COMMENT	Contact: Cordonnier-Pratt MM Laboratory for Genomics and Bioinformatics								

```

The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: PolyTMix
High quality sequence start: 30
High quality sequence stop: 638
POLYA=No.
Location/Qualifiers
1. .675
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone_lib="Dark Grown 1 (DG1)"
/note="Organ: 5-day-old dark-grown seedlings; Vector:
Lambda Zap; site 1: XhoI; Site 2: EcoRI; The library was
made from poly-A RNA in the cloning vector lambda Zap II.
Clones to be sequenced were prepared by mass excision."

ORIGIN
Alignment Scores:
Pred. No.: 2,16e-63 Length: 675
Score: 936.00 Matches: 178
Percent Similarity: 96.81% Conservative: 4
Best Local Similarity: 94.68% Mismatches: 6
Query Match: 51.26% Indels: 0
DB: 2 Gaps: 0

JIS-10-R29-432-4 (1-343) x BB357876 (1-675)

```

```

RESULT 5
BE355111
LOCUS
DEFINITION
  DG1_113_B07.g1_A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA
  sequence.
ACCESSION
  BE355111
VERSION
  1
KEYWORDS
  EST.
SOURCE
  Sorghum bicolor (sorghum)
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE
  1 (bases 1 to 691)
  Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and
  Pratt,L.H.
  An EST database from Sorghum: dark-grown seedlings
  Unpublished (2000)
  Contact: Cordonnier-Pratt MM
  Laboratory for Genomics and Bioinformatics
  The University of Georgia, Department of Plant Biology
  Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
  Tel: 706 542 1860
  Fax: 706 583 0210
  Email: mmpratt@uga.edu
  Sequences have been trimmed to exclude PolyA, vector and regions
  below Phred quality 16. The threshold for highest quality sequence
  is 20.
  Seq primer: PolyTMix
  High quality sequence start: 23
  High quality sequence stop: 682
  POLYA=No.
FEATURES
  source
  1..691
  /organism="Sorghum bicolor"
  /mol_type="mRNA"
  /db_xref="taxon:4558"
  /clone_lib="Dark Grown 1 (DG1)"
  /note="Organ: 5-day-old dark-grown seedlings; Vector:
  lambda zap; Site 1: XhoI; Site 2: EcoRI; The library was
  made from poly-A RNA in the cloning vector lambda ZAP II.
  Clones to be sequenced were prepared by mass excision."
ORIGIN
Alignment Scores:
Pred. No.: 1.36e-62 Length: 691
Score: 926.00 Matches: 176
Percent Similarity: 96.77% Conservative: 4
Best Local Similarity: 94.62% Mismatches: 6
Query Match: 50.71% Indels: 0
DB: 2 Gaps: 0
US-10-829-432-4 (1-343) x BE355111 (1-691)
QY 158 ArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrrleThrGlyLeuSerGlySer 177
Db 1 AGACAGAAAGTCTCGGCAAAAGCGTGTGTGGTATACAGGACTCAGTGGTTCA 60
QY 178 GlyLysSerThrLeuAlaCysAlaLeuSerArgGlyLeuHisCysArgGlyHisLeuThr 197
Db 61 GGGAAAGTACTTCTGTTGTGCTAGTCTAGTATGATGATGATATAGAGCCACATCAG 120
QY 198 TyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAla 217
Db 121 TATGACTTTGATGGTGCACAACTTAGACATGGCCCTCAATCGAGATTTAAGCTTTAAGGCA 180
QY 218 GluAspArgAlaGluAsnLeuArgValGlyGluValAlaLysLeuPheAlaAspAla 237
Db 181 GAAGACCGGTGCAGAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTGGCCATGCT 240
QY 238 GlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspAspAlaCysArg 257
  
```

```

Db 241 GGTATCATATGCAITTCAGTTCATATCTCCATACAGGAGAGATCGTATGTCATGCCGT 300
QY 258 AlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCys 277
Db 301 GCTCTACTTCAGATTTCTAACTTTATTTGAAGATTTATGGATTTGCCCTTAAAAATTTGT 360
QY 278 GluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPhe 297
Db 361 GAAGCTCGTATCTCTAAAGGTCGTGTAAGCTTGCACGCACAGGAAGATTAAGGTTTC 420
QY 298 ThrGlyIleAspAspProTyrGluProProLeuAsnGlyGluIleValIleLysMetLys 317
Db 421 ACTGGAATTTGATGATCCATACGACCGCCAGTTATTTGTTGATGATAGTAAATACGATGAA 480
QY 318 AspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGlu 337
Db 491 GATGGGGAATGCCCTTACCCAAAGCAATGGGAAGCAAGTCTTATCATACCTTGAAGAG 540
QY 338 AsnGlyTyrLeuGlnAla 343
Db 541 AACGGATATTGCAAGCT 558
RESULT 6
CK205911/c
LOCUS
DEFINITION
  FGAS017472 Triticum aestivum FGAS: Library 5 GATE 7 Triticum
  aestivum cDNA, mRNA sequence.
ACCESSION
  CK205911
VERSION
  1
KEYWORDS
  EST.
SOURCE
  Triticum aestivum (bread wheat)
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
  Pooideae; Triticeae; Triticum.
  1 (bases 1 to 1075)
  Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,
  Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroché,A.,
  Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilsson,D.,
  Penniket,C., Roach,J.L. and Sarhan,F.
  Functional Genomics of Abiotic Stress In Wheat and Canola Crops
  Unpublished (2003)
  Contact: Wm L Crosby
  Bioinformatics
  University of Saskatchewan, Department of Computer Science
  1C101 Engineering Building, 57 Campus Drive, Saskatoon,
  Saskatchewan, S7N 5A9, Canada
  Tel: 306 966 1769
  Fax: 306 966 2033
  Email: fgas_esta@usask.ca
  This sequence is the direct result of the Base calling software
  Phred (default parameters). It is the raw base calls. To aid in the
  identification of the high quality insert the software Lucy
  (default parameters) has been run on this sequence. Lucy identified
  the region [9,754].
  Plate: L5B004 row: J column: 15.
  Location/Qualifiers
  1..1075
  /organism="Triticum aestivum"
  /mol_type="mRNA"
  /db_xref="taxon:4565"
  /clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"
  /note="Vector: pCMV.SPORT6; Crown and developmental stages
  of spike formation in wheat cultivar Norstar. 4 mRNA
  populations were combined before constructing the library.
  The first mRNA population is from 1cm crown sections after
  30 days of cold acclimation. The second is from 1cm crown
  sections after 11 days of deacclimation (before
  deacclimation plants were fully vernalized for 49 days).
  The third is from different developmental stages of spike
  formation (5 to 50mm) that still have not emerged from the
  leaf (dissection required). The last is from different
  developmental stages of spike and seed formation after
  
```

having emerged from the leaf (visible) . First strand synthesis in this library was done in the presence of methylated dCTP thereby protecting from internal cleavage with NotI."

ORIGIN

Alignment Scores:			
Pred. No.:	7,69e-59	Length:	1075
Score:	891.50	Matches:	188
Percent Similarity:	78.2%	Conservative:	21
Best Local Similarity:	70.41%	Mismatches:	41
Query Match:	48.27%	Indels:	21
DB:	7	Gaps:	2
US-10-829-432-4 (1-343) x CK205911 (1-1075)			
Qy	77	ThrProArgSerSerHisSerSerAlaGlyLeuAlaSerAspSerGlyArgArgGluGly	96
Db	1073	AGCCCCCGTGGCGCTGGCGCTGGCGCTCCGTGGCCCACTTGGCCGGAACGGGGGTGC	1014
Qy	97	GluGlyArgGlyAlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgArg	116
Db	1013	CCGGGAACCGATGAACACCCCTC-----	987
Qy	117	ArgArgAsnGlyAlaAlaProGlyGluAlaProHisSerProValysGluIysProVal	136
Db	986	---AGGACGGGGCCCTTCG-----AAGCCCAAGTGGAAAAGCCTTTA	945
Qy	137	MetSerAsnIleGlyIysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSer	156
Db	944	AT-TCCAACTGGGAA-TCCAAATTAATTTTATGGCATGACTGCCCAATTTGGACAAAT	887
Qy	157	AspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTpolleThrGlyLeuSerGly	176
Db	886	GAGCAGACAGAAATGCTGGGGCCAAAGAGATTTTCATTGGATCAGAGACTCAGCGGT	827
Qy	177	SerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeu	196
Db	826	TCAGGAAAAGTACCTT-GCCTGTGCACCTGAT-CGGGGATTACACTACAGAGGCCACAC	769
Qy	197	ThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelys	216
Db	768	ACGTATGCTTGAATGGTGACAACTCAGACATGCGCTCAATCGAGATCTAAGCTTCAAG	709
Qy	217	AlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAsp	236
Db	708	GCAGAGACCGTAGAGAAATATACAGAGATTGGTGAAGTGGCCAAAGCTTTTCAGAT	649
Qy	237	AlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspAspAlaCys	256
Db	648	GCTGTACCATATGCATTGCTAGTTTGATATCTCCATACAGGAGAGACCGTGTATGATGC	589
Qy	257	ArgAlaLeuLeuProHisSerAsnPhelleGluValPheIleAspLeuProLeuIysIle	276
Db	588	CGAGCTCTACTTCCAGATTCTAGATTTATTTGAAGTATTTATGGATTTCACACTAGAATA	529
Qy	277	CysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGly	296
Db	528	TGTGAAGCTCGTGATCCTTAAGGGTTATACAAGCTTGCACGACAGGAAAGATTAAAGG	469
Qy	297	PheThrGlyIleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMet	316
Db	468	TTCCACGGAGTTGATGATCATCATCAAGAAATCCACCAAGTATGATGATGATTAAGATG	409
Qy	317	LysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCystYrLeuGlu	336
Db	408	GAAGGTGGGAATGCCCTTCACCGAAAGCAATGGCCCAAGTTCTATTCCTACCTTGAG	349
Qy	337	GluAsnGlyTyrLeuGlnAla	343
Db	348	AAAGATGGATATTTCAGGCT	328

CA204452	CA204452	614 bp	mRNA	linear	EST 25-SEP-2003
LOCUS	SCAGFL1086H06.g FL1 Saccharum officinarum cDNA clone SCAGFL1086H06				
DEFINITION	5', mRNA sequence.				
ACCESSION	CA204452				
VERSION	CA204452.1	GI:35241732			
KEYWORDS	EST.				
SOURCE	Saccharum officinarum				
ORGANISM	Saccharum officinarum				
REFERENCE	1 (bases 1 to 614)				
AUTHORS	Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.				
TITLE	The libraries that made SUCEST				
JOURNAL	Genet. Mol. Biol. 24 (1-4), 1-7 (2001)				
COMMENT	Contact: Arruda P Centro de Biologia Molecular e Engenharia Genetica Universidade Estadual de Campinas Caixa Postal 6010, 13083-970, Campinas SP, Brazil Tel: 55 19 3788 1137 Fax: 55 19 3788 1089 Email: parruda@unicamp.br Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bccccenter.fcav.unesp.br Plate: 086 Row: H Column: 06 Seq primer: T7 Promoter Primer. Location/Qualifiers 1. .614 /organism="Saccharum officinarum" /mol_type="mRNA" /db_xref="taxon:4547" /clone="SCAGFL1086H06" /lab_host="DH10B" /clone_lib="FL1" /note="Organ: Inflorescence at beginning of development (1cm-long); Vector: pSport1, Site_1: SalI; Site_2: NotI; (Inflorescence at beginning of development (1cm-long)). cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucest.lad.ic.unicamp.br/public"				
FEATURES	source				
ORIGIN					
Alignment Scores:					
Pred. No.:	1.38e-58	Length:	614		
Score:	874.00	Matches:	178		
Percent Similarity:	89.16%	Conservative:	3		
Best Local Similarity:	87.68%	Mismatches:	14		
Query Match:	47.86%	Indels:	8		
DB:	6	Gaps:	3		
US-10-829-432-4 (1-343) x CA204452 (1-614)					
Qy	45	AlaValIleLeuValAsnProGlnArgAlaProProValLeu-ProGlyLeuThrProSe	64		
Db	18	GCCGTGCGCCCGCTAAATCCATAGCGCGCGGTCCGTCTCCCTCCCTCC	71		
Qy	64	rAspAlaProLeuProAlaLeuValIleHisGlyLeuThrProArgSerHisSerSe	84		
Db	72	CGATGCTCCGTGCGCGCTCGTGTATCCATGTCCTCCCTCCCTCCCTCCCT	131		
Qy	84	rAlaGlyLeuAlaSerAspSerGlyArgArgGluGlyGlu-----GlyArgGlyAl	101		
Db	132	CGCGGGCTCCGACGCGGTGTGGCGCGCGAGGGGGAGGAGGAGGCGCGGTGC	191		
Qy	101	aArgThrHisCysHisArgGlyIleGlyArgTrpValArgArgArgArgAsnGlyAl	121		

```
Db 192 GCGTGGCACTGCCCGCGCATTCGGCGGTGGATG---CGGTGGCGCGGAATGGAGC 248
Qy 121 aalaPro-GlyGluAlaProHisSerProVallyGluLysProValMetSerAenileG 141
Db 249 AGCGCCACGGGAGGACACACAGCGGCCAGTGAAGAGAGCCTGTAATGTCAACATTTG 308
Qy 141 lYlYSerThrAenileLeuTPhHisAsnCysLeuileGlyGlnSerAspArgGlnLysL 161
Db 309 GGAATCAACTAATATTTATGGCATAATTTGCCGATTTGGCAATCTGATAGACAGAAGT 368
Qy 161 euLeuGlyClnLysGlyCysValValTrrPilleThrGlyLeuSerGlySerGlyLysSerT 181
Db 369 TGCTGGGACAAAAGCGCTGTGCTGTGGATAACAGGACTCAGTGGTTCAGGGAAGTA 428
Qy 181 hcLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuA 201
Db 429 CTCCTTGCATGTGCTGAGTGTGAGTGTGATATATAGAGCCACCTTCACATATGTACTTG 488
Qy 201 spGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelysAlaGluAspArgA 221
Db 489 ATGTTGACAACTTAGACATGCGCTAAATCGAGATTAAAGCTTTAAGCGCAGAAGCCGTG 548
Qy 221 lsGluAsnileArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValileC 241
Db 549 CAGAAAATATACGAAGAGTTGTTGAAGTGGCAAGCTTTTTCGCCGATGCTGTATCATAT 608
Qy 241 yaile 242
Db 609 GCATT 613
```

```
RESULT 8
CD863007 617 bp mRNA linear EST 11-JUL-2003
LOCUS AZO1.105F21F010130 AZO1 Triticum aestivum cDNA clone AZO1105F21,
DEFINITION mRNA sequence.
ACCESSION CD863007
VERSION CD863007.1 GI:32546823
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 617)
AUTHORS Genoplante.
TITLE Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr/).
```

```
FEATURES
source
1..617
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="AZO1105F21"
/tissue_type="leaf"
/clone_lib="AZO1"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 7,07e-58 Length: 617
Score: 865.00 Matches: 163
Percent Similarity: 93.51% Conservative: 10
Best Local Similarity: 88.11% Mismatches: 12
Query Match: 47.37% Indels: 0
```

```
DB: 6 Gaps: 0
US-10-829-432-4 (1-343) x CD863007 (1-617)
Qy 159 GlnLysLeuLeuGlyGlnLysGlyCysValValTrrPilleThrGlyLeuSerGlySerGly 178
Db 1 CAGAAAATGCTGGGGCAAAAAGAGTGTGTATATGGATCACAGGAGCTCAGCGGTTTCAGGG 60
Qy 179 lYsSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyr 198
Db 61 AANAAGTACCGTGTCTGTGCACTGAGTCGGGAATTTACACTACAGAGGCCACACACGATAT 120
Qy 199 ValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelysAlaGlu 218
Db 121 GTTCTTGTATGTCGACAACTCAGACATGGCTTAATCGAGATCTTAAGCTTCAAGGCAGAA 180
Qy 219 AspArgAlaGluAsnileArgArgValGlyGlnValAlaLysLeuPheAlaAspAlaGly 238
Db 181 GACCGTACAGAAAATATACGAAGAGTTGGAGAAGTGGCAAGCTTTTTCGAGATGCTGGT 240
Qy 239 VallileCysilleAlaSerLeuileSerProTyrArgArgAspArgAlaCysArgAla 258
Db 241 ACCATATGCATGCTAGTTGTATCTCCATACAGGAGAGCCGTGATGTCATGCCAGCT 300
Qy 259 LeuLeuProHisSerAsnPhelleGluValPheileAspLeuProLeuLysileCysGlu 278
Db 301 CTACTTCCAGATTCTAGATTTTATGAGTATTTATGGATTTTCCACCTAGATAATATGTCAA 360
Qy 279 AlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysileGlyPheThr 298
Db 361 GCTCGTATCTCTAAAGGGTTATACAAGCTTTCACGCA CAGGAAAGATTAAAGGGTTTACC 420
Qy 299 GlylleAspAspProTyrGluProPheleAsnGlyGluileValileLysMetLysAsp 318
Db 421 GGAGTGTATGATTCATACGATCACCAGTGAATAGTAGATAGTAGTAATTAAGATGAAGGT 480
Qy 319 GluGluCysProSerProLysAlaMetAlaLysGlnValLeuLysCysTyrLeuGluGluAsn 338
Db 481 GGGGAATGCCCTTCAACGAGGCAATGCCCGCAGCAAGTTCTGTCTCTACCTTGAGAAGAAC 540
Qy 339 GlyTyrLeuGlnAla 343
Db 541 GGATATTTTCAGGCT 555
```

RESULT 9

```
CD8285854 860 bp mRNA linear EST 27-FEB-2004
LOCUS CR285854 Oryza sativa library (Han B) Oryza sativa cDNA clone
DEFINITION y625e05p5, mRNA sequence.
ACCESSION CR285854
VERSION CR285854.1 GI:44672420
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 860)
AUTHORS Han, B., Feng, Q., Huang, Y. C., Ying, K., Li, Y., Guan, J. P., Zhu, J. J.,
Zhao, Q., Hu, X., Liu, Y. L., Mu, J., Yu, Z., Chen, L., Fan, D. L.,
Weng, Q. J., Zhang, L., Lu, Y. Q., Yu, S. L., Liu, X. H., Lu, T. T.,
Zhang, Y. J., Lu, Y., Li, C., Li, T., Zhang, Y., Hu, H., Jia, P. X.,
Zhang, L., Lan, L. F., Chen, W., Wu, S. A. and Xue, Y. B.
Rice cDNA EST clone
Unpublished (2003)
Contact: Han Bin
National Center for Gene Research
Chinese Academy of Sciences
500# Cao Bao Road, Shanghai 200233, China
Email: bhan@ncgr.ac.cn
Clone requests: bhan@ncgr.ac.cn
This is rice cDNA est clone
Web site: http://www.ncgr.ac.cn.
```

TITLE
JOURNAL
COMMENT

National Center for Gene Research
Chinese Academy of Sciences
500# Cao Bao Road, Shanghai 200233, China
Email: bhan@ncgr.ac.cn
Clone requests: bhan@ncgr.ac.cn
This is rice cDNA est clone
Web site: <http://www.ncgr.ac.cn>.


```
FEATURES
  source
    Location/Qualifiers
      1..860
      /organism="Oryza sativa"
      /mol_type="mRNA"
      /db_xref="taxon:4530"
      /clone="y625e05p5"
      /clone_lib="Oryza sativa library (Han B)"

ORIGIN
Alignment Scores:
  Pred. No.: 1.46e-57 Length: 860
  Score: 863.50 Matches: 166
  Percent Similarity: 85.44% Conservative: 10
  Best Local Similarity: 80.58% Mismatches: 25
  Query Match: 47.29% Indels: 5
  DB: 7 Gaps: 1

US-10-829-432-4 (1-343) x CR285854 (1-860)
QY 106 HisArgGlyIleGlyArgTrpValArgArgArgAsnGlnGlyAlaProGlyGlu 125
DB ::::|||||
QY 126 AlaProHisSerProValysGluLysProValMetSerAsnIleGlyLysSerThrAsn 145
DB ::::|||||
QY 146 IleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuGlyGlnLys 165
DB ::::|||||
QY 166 GlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAlaCysAla 185
DB ::::|||||
QY 186 LeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpValLeuAspIlyAspAsnLeu 205
DB ::::|||||
QY 206 ArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnLeuArg 225
DB ::::|||||
QY 226 ArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeu 245
DB ::::|||||
QY 246 IleSerProTrpArgArgAspArgAlaCysArgAlaLeuLeuProHisSerAsnPhe 265
DB ::::|||||
QY 266 IleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProLysGlyLeu 285
DB ::::|||||
QY 286 TyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAspProTyrGlu 305
DB ::::|||||
QY 306 ProProIleAsnGlyGlu 311
DB ::::|||||
QY 311 TCAACAGTACTAGTGCAG 619
DB ::::|||||

RESULT 10
CAL182186 673 bp mRNA linear EST 24-SEP-2003
LOCUS SCBGST3108C02.g ST3 Saccharum officinarum cDNA clone SCBGST3108C02
5', mRNA sequence.
ACCESSION CAL182186
VERSION CAL182186.1 GI:35117704
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
```

REFERENCE

AUTHORS Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
TITLE The libraries that made SUCEST
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: paruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 108 row: C column: 02
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1..673
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCBGST3108C02"
/lab_host="DH10B"
/clone_lib="ST3"
note="Organ: Fourth apical stalk internodes of adult
plants; Vector: pSPort1; Site_1: SalI; Site_2: NotI; An
unidirectional cDNA library generated from [Fourth apical
stalk internodes of adult plants]. cDNA was prepared from
polyA+ mRNA using SuperScript plasmid System kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

FEATURES

source
1..673
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCBGST3108C02"
/lab_host="DH10B"
/clone_lib="ST3"
note="Organ: Fourth apical stalk internodes of adult
plants; Vector: pSPort1; Site_1: SalI; Site_2: NotI; An
unidirectional cDNA library generated from [Fourth apical
stalk internodes of adult plants]. cDNA was prepared from
polyA+ mRNA using SuperScript plasmid System kit
(Invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Alignment Scores:
Pred. No.: 1.14e-55 Length: 673
Score: 837.50 Matches: 186
Percent Similarity: 75.30% Conservative: 3
Best Local Similarity: 74.10% Mismatches: 27
Query Match: 45.87% Indels: 36
DB: 6 Gaps: 5
US-10-829-432-4 (1-343) x CA182186 (1-673)
QY 16 ThrGlnGlnProProSerProAlaProGlyProAlaSerGlnGlnArgGlnGlyAsn 35
DB ::::|||||
DB 12 AGCCAGCAACACCTCTTCGGTCCGACCTGACCGCGGAGCGGCCACCGC----- 62
QY 36 ThrLeuLeuSerProThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaPro 55
DB ::::|||||
DB 63 -----GCCGTCGGCCCGTAAATCCATAGCGCGCGCT 95
QY 56 ProValLeuProGlyLeuThr-ProSerAspAlaProLeuProAlaLeuValIleHisG1 75
DB ::::|||||
DB 96 CCGGTCCTCCCGCCACCCCGCCGCTGCTGGAGCTCGTGATCCAT-- 153
QY 75 yLeuThrProArgSerSerHisSerAlaGlyLeuAlaSerAspSerGlyArgArgG1 95
DB ::::|||||
DB 154 -----GGCCTCCGACGGGTGGTGGCGCGCGA 182
QY 95 uGly-----GluGlyArgGlyAlaArgThrHisCysHisArgGlyIleGlyArgTr 112
DB ::::|||||
DB 183 GGGGAGGAGAGGAGGCGCGTGGCGTGGCGACTGCCACCGCGCATTTGGGCGGTG 242
QY 112 pValArgArgArgArgAsnGlyAlaAlaPro-GlyGluAlaProHisSerProValL 132
DB ::::|||||

Db 243 G-----CGCGGGAATGAGCAGCGCCACGGGAG-----GGCCCAAGTGA 281

Qy 132 ysGluysProValMetSerAnilleGlyLysSerThrAsnIleLeuTrpHisAsnCysL 152

Db 282 AGGAGAGCGCTGATGTCACACATGGGAATCAACTAATATTTATGGCATAATGCC 341

Qy 152 euileGlycInSerAspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleT 172

Db 342 CGAATGGGACAACTCTGATCGACAGAAGTTGCTGGGACAAAAGGCTT-GTCGTGTGGATA 400

Qy 172 hrGlyLeuSerGlySerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisC 192

Db 401 CAGGACTCAGTGGTTCAGGGAAGTACTCTTGCATGTGCACTGAGTCATGAGTTGCATT 460

Qy 192 ysArgGlyHisLeuThrTrValLeuAspGlyAspAsnLeuArgHisGlyLysLeuAsnArg 212

Db 461 ATAGAGGCCACCTCACATATGTACTTGTGATGTGACAACTTAGACATGGCCTAATCGAG 520

Qy 212 spLeuSerPhelysAlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaL 232

Db 521 ATTTAAGTTTAAAGCAGAGACCGTGCAGAAAATATACGACGAGTTGGTGAAGTGGCAA 580

Qy 232 ysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyArgArgA 252

Db 581 AGCTTTTTCGCCATGCTGATCATATGCAATTCATTTGCTAGCTTGATATCTCCATACAGGAG 640

Qy 252 spArgAspAlaCysArgAlaLeuLeuPro 261

Db 641 ATCGTGTGATGTCGCGGGCTCTACTTCCA 669

RESULT 11
LOCUS CO095086 867 bp mRNA linear EST 16-JUN-2004
DEFINITION GR_Eal7119.r GR_Ea Gossypium raimondii cDNA clone GR_Eal7119 3', mRNA sequence.

ACCESSION CO095086

VERSION CO095086.1 GI:48793772

KEYWORDS EST.

SOURCE Gossypium raimondii

ORGANISM Gossypium raimondii

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

1 (bases 1 to 867)

Kim H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,

Udall, J.A., Rapp, R.A., Wendel, J.F., Rao, K., Soderlund, C. and

Wing, R.A.

Global assembly of Cotton ESTs

Unpublished (2004)

Contact: Rod A. Wing

Arizona Genomics Institute

The University of Arizona

Forbes Building Room 303, Tucson, AZ, 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: <http://genome.arizona.edu>

Plate: 17 row: I column: 19.

Location/Qualifiers

1..867

/organism="Gossypium raimondii"

/mol_type="mRNA"

/db_xref="taxon:29730"

/clone="GR_Eal7119"

/tissue_type="whole seedlings"

/dev_stage="first true leaves"

/clone_lib="DH10B"

/clone_lib="GR_Ea"

/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:

EcoRV; library made by Invitrogen with RNA supplied by

Wendle lab. Directional cloned into NotI-EV. Colonies

plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Alignment Scores:

Pred. No.: 7.17e-54 Length: 867

Score: 816.50 Matches: 152

Percent Similarity: 85.44% Conservative: 24

Best Local Similarity: 73.79% Mismatches: 29

Query Match: 44.72% Indels: 1

DB: 7 Gaps: 1

US-10-829-432-4 (1-343) x CO095086 (1-867)

Qy 137 MetSerAnilleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSer 156

Db 45 ATGTCGACTATGGGAAATTCGACAAATATATTTTGGCAAGAAATCCCTGTGGGAAGCTT 104

Qy 157 AspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGly 176

Db 105 GAGAGCGGAGCTACTTACCAAAAGGGTTGTGTATGGATCCTGGTCTCAGCGGA 164

Qy 177 SerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeu 196

Db 165 TCAGGTAAAGCACACTTGCAATGTTCTTAAGTTGGGAACCTTCATACAAGAGGCAAGCTA 224

Qy 197 ThrTrpValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhelys 216

Db 225 TCTTACATACTTGTATGGGACACCGTTCCGCCAGGATTTAAACAAGAGATCTTGGTTTCAAG 284

Qy 217 AlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAsp 236

Db 285 GCTGAGGATCGACAGAAAATATTCGCAGGGTTGGTGNAGTGGCAAGCTCTTTCAGAT 344

Qy 237 AlaGlyValIleCysIleAlaSerLeuIleSerProTyArgArgAspAspAlaCys 256

Db 345 GCTGGTTTAACTGCAATTCGCCAGTAAATATCTCCATATGGGAAAGACCGTGATCATGC 404

Qy 257 ArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIle 276

Db 405 CGAGCAATGTTGCCGGATGCT---TTTATTGAGGTTTTCATGAACATCCCTAGCATTA 461

Qy 277 CysGluAlaArgAspProLysGlyLeuTyLysLeuAlaArgThrGlyLysIleLysGly 296

Db 462 TGTGAGAGCGCAGATCCGAAAGGTTCTTACAAAGCTTGCCCGTGTGGAAAGATTAAAGT 521

Qy 297 PheThrGlyIleAspAspProTyArgProPheLeuGluIleValIleLysMet 316

Db 522 TTTACTGCAATAGATGATCCTTACGACCACTTTTGAACCTGTGAGATAGAACTAATAATCAG 581

Qy 317 LysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyLeuGlu 336

Db 582 AAAGATGGAGTTTGTCCACACACCTAGTCCCATGCTGGGGAAGTAATTAATTACTTCTGGAG 641

Qy 337 GluAsnGlyTyLeuGln 342

Db 642 GACAAAGGATATCTGCAG 659

RESULT 12

LOCUS CNS0A87V

DEFINITION

CNS0A87V 1066 bp mRNA linear HTC 06-FEB-2004

Arabidopsis thaliana Full-length cDNA Complete sequence from clone

GSU3138ZF03 of Silique of strain col-0 of Arabidopsis thaliana

(thale cress).

ACCESSION BX821394

VERSION BX821394.1 GI:42467158

KEYWORDS HTC; GSLT cDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1066)

Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,

Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,

Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.

Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:

TITLE

A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
 Unpublished
 2 (bases 1 to 1066)
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Direct Submission
 Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 The sequences are based on single pass reads.
 Life Technologies (a division of Invitrogen) members carried out
 full-length libraries construction : Temple G.
 Genoscope members carried out sequencing and annotation : Castelli
 V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
 Schachter V., Weissenbach J., Salanoubat M.
 URGV INRA : Clepet C., Caboche M.
 Annotation is based on the June 2003 version of the Arabidopsis
 genome released by MIPS (Munich Information center for Protein
 Sequences). 5 prime and 3 prime are assembled with Phrap.
 http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
 length
 http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
 FEATURES
 source
 1..1066
 /location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Col-0"
 /db_xref="taxon:3702"
 /clone="GSUFSIU382F03"
 /tissue_type="Silique"
 /plasmid="PCMVSPORT_6"
 1..1066
 /gene="At2g14750"

ORIGIN

Alignment Scores:
 Pred. No.: 1.04e-53 Length: 1066
 Score: 816.00 Matches: 159
 Percent Similarity: 77.8% Conservative: 31
 Best Local Similarity: 65.1% Mismatches: 44
 Query Match: 44.6% Indels: 10
 DB: 3 Gaps: 3

US-10-829-432-4 (1-343) x CNS0A87V (1-1066)

Qy 100 GlyAlaArgThrHisCysHisArgGlyLeGlyArgTTPValArgArgArgAsn 119
 Db 204 GGATCTCAAACTCTGAGTCATAACAAAATGGATCTATTCTGAGGTTAAATCCATTAAAC 263
 Qy 120 GlyAlaAlaProGlyGluAlaProHisSerProValHisGluLysProValMetSerAsn 139
 Db 264 GGT-----CACACGGGACAAAGCAAGGACCT---TTGTCTACG 299
 Qy 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuLeGlyGlnSerAspArgGln 159
 Db 300 GTCCGAAACTCCACAAATATAAAGTGGCATGAATCTTCTGTGAGAAAGTTGATAGACAG 359
 Qy 160 LysLeuLeuGlyGlnLysGlyCysValValTTPleThrGlyLeuSerGlySerClyLys 179
 Db 360 AGATTGCTTGATCAGAAAGGAGTGTGATTGGTCACCGGCTTTAGTGGTTCCAGGGAAG 419
 Qy 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 199
 Db 420 AGTACTTTGGCTTGCTTTGAATCAGATGTTGTATCAAAAGGGAAGCTTTGTATTATT 479
 Qy 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
 Db 480 CTTGATGGTGATAATGTTAGGCATGGCTTAAACCGGTGATCTTAGCTTTAAAGCTGAGGAT 539
 Qy 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
 Db 540 CGTCAGAGATATTTCGTAGAGTTGGAGAGGTTGCTTAAGCTTTTTCGGAGTGTGGAATA 599

Qy 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspAlaCysArgAlaLeu 259
 Db 600 ATCTGCATTGGAGCTTTGATATCTCTATAGAACAGATAGGACGCTTGTGGAAGTTTG 659
 Qy 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
 Db 660 CTCGCCGAGGAGATTTTGTGAGGTCTTACAGCTCTCGTCAGAGAAAGATCAAAAGGTTTACCCGGG 719
 Qy 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
 Db 720 AGGATCAAAAGGCTTTTACAGCTCTCGTCAGAGAAAGATCAAAAGGTTTACCCGGG 779
 Qy 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
 Db 780 ATCGATGACCTTAGGACCACTTGAACCTCCGAG-----ATTCTCTAGGACGTGAA 833
 Qy 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
 Db 834 GGAGCAACTTCTCCTATCGAAATGCGGAAAGGTCGTTCGATATCTTAGATAACAAGGCT 893
 Qy 340 TyrLeuGlnAla 343
 Db 894 TATCTTCAAGCA 905

RESULT 13
 AY110087
 LOCUS Zea mays CL1019_1 mRNA sequence. 991 bp mRNA linear HTC 17-OCT-2002
 DEFINITION Zea mays CL1019_1 mRNA sequence.
 ACCESSION AY110087
 VERSION AY110087.1 GI:21214175
 KEYWORDS HTC.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 991)
 AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
 Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
 TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
 Overgo Probes
 JOURNAL Unpublished (2002)
 REFERENCE 2 (bases 1 to 991)
 AUTHORS Coe, E.H.
 TITLE Direct Submission
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA
 COMMENT If you are interested in getting corresponding physical clones,
 these are publicly available from ZmDB and may be found by BLAST
 searching at MSI, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
 www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat
 Schnable, Iowa State, then clones may be requested from ZmDB:
 www.zmdb.iastate.edu.
 FEATURES
 source
 1..991
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="maizdb:629888"
 /db_xref="taxon:4577"
 /clone_lib="Maize Mapping Project/DuPont Consensus
 Library"
 /note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 contigs to seed DuPont contigs; this resource was
 assembled by DuPont as part of a collaboration for the
 overgo addressing of BACs in conjunction with the Maize
 Mapping Project"

ORIGIN
 Alignment Scores: 2.54e-53 Length: 991
 Pred. No.: 810.50 Matches: 162
 Score:

Percent Similarity:	74.06%	Conservative:	15
Best Local Similarity:	67.78%	Mismatches:	24
Query Match:	44.39%	Indels:	39
DB:	3	Gaps:	1
US-10-829-432-4 (1-343) x AY110087 (1-991)			
QY	136	ValMetSerAsnIleGlySerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGln	155
DB	58	GTGACCTCGACCGTCGGGAAATCGACGAATCTCTGGCATGAGTGGCCATCGGGCAG	117
QY	156	SerAspArgGlnLysLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSer	175
DB	118	AGGAGCGACAGGCTCTGCTGAACCAAGAGGGCTGCTGTGGATCATCGCCCTAAGC	177
QY	176	GlySer-----	177
DB	178	GGTTC-AGGTGAGGCTGCAAGAGCTGATCTTTCATTTTAATTACGAACAAGATCCT	236
QY	177	-----	177
DB	237	TCCTTTTATATATATACGTAGGAATTCGGCAGGAGGATCACTGGCCCTGACTTCATCA	296
QY	178	GlyLysSerThrIleuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThr	197
DB	297	GGGAAAGACGCGTCCGCTGCGCGCTGAGCGCGAGCTGCACGGCAGAGGCCACCTCAGC	356
QY	198	TyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAla	217
DB	357	TAGGTCCTCGACGGCGACAACTCTAGGCACGGGCTGAAACAGGAGCTCAGCTTCGGAGCA	416
QY	218	GluAspArgAlaGluAsnIleArgValGlyGluValAlaLysLeuPheAlaAspAla	237
DB	417	GAGGACCGCGCGAGAACATCCGACAGTAGGGGAAGTAGCGAGCTGTTCGCCACGCT	476
QY	238	GlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspAspAlaCysArg	257
DB	477	GGCCTCGTCTGCATCGCCAGCCCTCATATCGCCCTACAGAAGCGCGGTGTGCG	536
QY	258	AlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCys	277
DB	537	GATCTGCTGCCAACGACACTCTTTATCGAGGTGTTCCTGGACGTCCGGCTTCAAGTGTC	596
QY	278	GluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPhe	297
DB	597	GAGCCAGGAGCCCAAGGCTCTACAGCTCGCAGCGCGCGGCAATCAAGGGTTC	656
QY	298	ThrGlyIleAspAspProTyrGluProPheIleAsnGlyGluIleValIleLysMetLys	317
DB	657	ACCGGCATCGACGATCCTTACGAACCGCGCTCGGACTCGGAGATGATGATCCAGTGTRAA	716
QY	318	AspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGlu	336
DB	717	GTCGGGCACTGCCCTTCGCTCGCTGAATCGATGGTGTGTCACGTGTGTGTCGTTGAG	773
RESULT 14			
CA900730			
LOCUS			
DEFINITION			
PCSC14960 Scarlet Runner Bean Suspensor Region Triplex2 Phaseolus			
coccineus cDNA 5', similar to Putative adenylyl sulfate kinase, mRNA			
sequence.			
CA900730			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Phaseolus coccineus			
Phaseolus coccineus			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;			
Phaseolus.			
1 (bases 1 to 659)			
REFERENCE			
AUTHORS			
Bui,A.Q., Le,B.H., Weterings,K., Bi,Y.-P., Choi,J.-S.,			
McElroy,K.E., Choi,P.S., Harada,J.J., Fischer,R.L. and			

Goldberg,R.B.
Gene Activity in Different Regions of a Post-Fertilization Plant
Embryo by EST Analysis
Unpublished (2002)
Contact: Goldberg, R.B.
Department of Molecular, Cell, & Developmental Biology
University of California, Los Angeles
621 Charles E. Young Drive South, Los Angeles, CA 90095-1606, USA
Tel: 310 825 3270
Fax: 310 825 8201
Email: bobg@ucla.edu
Seq primer: 5' Triplex
POLYA=No.

FEATURES
Source
1..659
/organism="Phaseolus coccineus"
/mol_type="mRNA"
/cultivar="Hammond's Dwarf Scarlet"
/db_xref="taxon:3886"
/dev_stage="6-days post-pollination"
/clone_lib="Scarlet Runner Bean Suspensor Region Triplex2"
/note="Organ: Suspensor Region of Globular-Stage Embryos;
vector: Triplex2; Site 1: SfIIA; Site 2: SfIIB; Suspensor;
regions were micro-dissected from globular-stage embryos
six days after pollination from greenhouse-grown plants
[Weterings et al., Plant Cell 13, 2409-2425 (2001)].
Double-stranded cDNA was synthesized from suspensor mRNA
using the SMART cDNA Library Construction Kit according to
the manufacturer (Clontech). The suspensor cDNA fragments
were directionally ligated into the SfII restriction site
of the lambda Triplex2 vector (Clontech), and the
recombinant cDNAs were transformed into E. coli XL1-Blue
cells (Clontech). Suspensor cDNA plasmids used for
directional sequencing were obtained by in vivo excision
from the lambda Triplex2 recombinants in E. coli BM25.8
cells (Clontech)."

ORIGIN
Alignment Scores:
Pred. No.: 7,396-53 Length: 659
Score: 801.50 Matches: 149
Percent Similarity: 85.50% Conservative: 22
Best Local Similarity: 74.50% Mismatches: 28
Query Match: 43.89% Indels: 1
DB: 6 Gaps: 1
US-10-829-432-4 (1-343) x CA900730 (1-659)
QY 137 MetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSer 156
DB 60 ATGGCAACTCTGAGCAATTCACCAATATCTTTGGCAAGATTGTCAATAGGAAGGCCT 119
QY 157 AspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGly 176
DB 120 GAAAGGCAGCAGCTACTTAACCAAAAGCGATGTGTGTATGATGATTCTGGAATCGACGGA 179
QY 177 SerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeu 196
DB 180 TCAGAAAAGACACATTCGGCATGTTCCTTAAGCAGAGAACTGTATCAAGGGGAAAGTTA 239
QY 197 ThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLys 216
DB 240 TCTTATGCTTGTATGGAGATTAACCTTAGGATGAGCTTAACACAGAGATCTTGTTTAAA 299
QY 217 AlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAsp 236
DB 300 ACTGACGACCGCAGTGAACCAATTCGAGAACTCGAGAAAGTGGCAAACTCTTTCTGAT 359
QY 237 AlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspAspAlaCys 256
DB 360 GCGGGTTTGATATGTGTGCCAGTCTGATATCTCTTACAGAAGAGATCGAGACATTCG 419
QY 257 ArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIle 276

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2005, 11:36:00 ; Search time 120 seconds
(without alignments)
1463.693 Million cell updates/sec

Title: US-10-829-432-4
Perfect score: 1826
Sequence: 1 RPFHFINTQTEPLVTHTHQPP.....PRMAKQVLCYLENGYLQA 343

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	873.5	47.8	345	2 Q6ZL22	Q6ZL22 oryza sativ
2	834	45.7	208	2 Q9SRW7	Q9SRW7 arabidopsis
3	833	45.6	312	1 KAP1_CATRO	Q49204 catharanthu
4	816	44.7	276	1 KAP1_ARATH	Q43295 arabidopsis
5	816	44.7	276	2 Q8LES2	Q8LES2 arabidopsis
6	812.5	44.5	288	2 Q9SE92	Q9SE92 zea mays (m
7	777.5	42.6	293	1 KAP2_ARATH	Q49196 arabidopsis
8	757.5	41.5	310	2 Q84JF0	Q84JF0 arabidopsis
9	755	41.3	290	2 Q9FJX1	Q9FJX1 arabidopsis
10	752.5	41.2	305	2 Q8LP64	Q8LP64 arabidopsis
11	632	34.6	152	2 Q9SQR9	Q9SQR9 arabidopsis
12	571	31.3	200	1 CYSC_CLOAB	Q97mt8 clostridium
13	563	30.8	618	2 Q9NDP8	Q9NDP8 ciona intes
14	555	30.4	199	1 CYC2_BACSU	Q06735 bacillus su
15	555	30.4	230	2 Q73WR1	Q73WR1 mycobacteri
16	550.5	30.1	201	2 Q8DGK5	Q8DGK5 synechococc
17	543.5	29.8	199	2 Q8Q3N3	Q8Q3N3 oceanobacill
18	538.5	29.5	207	1 CYSC_LACPL	Q88x60 lactobacill
19	534.5	29.3	202	1 CYC1_BACHD	Q9Kct0 bacillus ha
20	528.5	28.9	233	2 Q7XBB6	Q7XBB6 porphyra pu
21	527.5	28.9	614	2 Q842M0	Q842M0 rhodococcus
22	524.5	28.7	199	2 Q8CR04	Q8CR04 staphylococ
23	524	28.7	271	2 Q81351	Q81351 enteromorph
24	522.5	28.6	208	2 Q6LM71	Q6LM71 photobacter
25	521.5	28.6	197	2 Q63DV7	Q63DV7 bacillus ce
26	520.5	28.5	197	2 Q73B74	Q73B74 bacillus ce
27	519.5	28.5	197	2 Q6HLD2	Q6HLD2 bacillus th
28	519.5	28.5	208	2 Q7V3B6	Q7V3B6 prochloroc
29	519	28.4	205	2 Q6DIA7	Q6DIA7 erwinia car
30	514.5	28.2	197	2 Q81T47	Q81T47 bacillus an
31	514.5	28.2	214	2 Q7U3Z4	Q7U3Z4 synechococc

RESULT 1

ID	Q6ZL22	PRELIMINARY;	PRT;	345 AA.
AC	Q6ZL22;			
DT	05-JUL-2004 (Tremblrel. 27, Created)			
DT	05-JUL-2004 (Tremblrel. 27, Last sequence update)			
DT	05-JUL-2004 (Tremblrel. 27, Last annotation update)			
DE	Putative adenosine-5'-phosphosulfate kinase.			
GN	Name=OJ1699_E05.15;			
OS	Oryza sativa (japonica cultivar-group).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzeae; Oryza.			
OX	NCBI_TaxID=39947;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Sasaki T., Matsumoto T., Yamamoto K.;			
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AP003845; BAC83149.1; -			
DR	GO; GO:0005524; P:ATP binding; IEA.			
DR	GO; GO:0016301; P:kinase activity; IEA.			
DR	GO; GO:0006772; P:transferase activity; IEA.			
DR	GO; GO:000103; P:sulfate assimilation; IEA.			
DR	InterPro; IPR002891; APS_kinase.			
DR	Pfam; PF01583; APS_kinase; 1.			
DR	ProDom; PD002350; APS_kinase; 1.			
DR	TIGRFAMs; TIGR00455; _apsK; 1.			
KW	Kinase.			
SQ	SEQUENCE 345 AA; 37748 MW; BE4235F256676698 CRC64;			
Query Match				
Best Local Similarity 53.1%; Score 873.5; DB 2; Length 345;				
Matches 188; Conservative 33; Mismatches 68; Indels 65; Gaps 7;				
QY	10 EPLVTHQPPSPAP----	GPASOG-----	ORQNTLLSPPTTAVILVNP----	QRAPP 56
Db				
Db	33 DPRATARVVPAAAPVRSRSPANLGLP	PHPPRRLRLAPPRITAAVTGGPRRRRAPP	92	
QY	57 VLPGLTPSDALPALVIHGLTPRSSH	SAGLASDGR-----	EGEGRGARTHCHRG 110	
Db				
QY	93 PL-----	ECAGSSSSSLRRPREEESEESSTAHAGV-	127	
Db				
QY	111 RWVRRRRNGAAPGEAPHSPVKEKPV	M-----	SNIGKSTNILWHNCLIGSQDQKLGQKGC	167
Db				
QY	168 VVWITGLSGSKSTLACALSRELHCR	GLHTVYVLDGDNLRHGLNRDLSKFAEDRAENIRRV	227	
Db				
QY	170 VVWITGLSGSKSTLACTDLRELH	TRGKLSVYVLDGDNLRHGLNKDGLGKFAEDRAENIRRV	229	
Db				
QY	228 GEVAKLFADAGVICIASLISPYR	DRDRACRALLPHSNFIEVFDLPLKICEARDPKGLYK	287	
Db				
QY	230 GEVAKLFADAGLVCIASFISPYR	DRDRSCRLSDGGSFIEVFLNMLPLCESRDPKGLYK	289	
Db				

Q92vh6 rhizobium m
Q81fy9 bacillus ce
Q82bp3 versinia pe
Q66ec6 versinia ps
Q90xy2 fugu rubrip
Q9A882 caulobacter
Q12657 penicillium
Q22501 caenorhabdi
Q93hs6 bradyrhizob
Q9p7g9 schizosacch
Q9eyyl klebsiella
Q27128 u bifunctio
Q07309 r nodg bifu
Q7n8k9 photorhabdu

```

OC lamids; Gentianales; Apocynaceae; Rauvolfioideae; Vinceae;
OC Catharanthus.
OC NCB1_TaxID=4058;
RN [1]
RP SEQUENCE FROM N.A.
RA Schiffmann S., Schwenn J.-D.;
RT "Isolation of cDNA clones encoding adenosine-5'-phosphosulfate-kinase
RT (SC2.7.1.25) from Catharanthus roseus and an isoform (akn2) from
RT Arabidopsis."
RL (er) Plant Gene Register PGR98-116.
CC -1- FUNCTION: Catalyzes the synthesis of activated sulfate.
CC -1- CATALYTIC ACTIVITY: ATP + adenylyl sulfate = ADP + 3'-
CC phosphoadenylyl sulfate.
CC -1- PATHWAY: Sulfate activation; cysteine biosynthesis reductive
CC branch; second step.
CC -1- SUBCELLULAR LOCATION: Chloroplast (By similarity).
CC -1- SIMILARITY: Belongs to the APS kinase family.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF044285; AAC31145.1; -.
CC PIR: T08076; T08076.
CC DR HSP; Q12657; I7M7G.
CC DR InterPro; IPR002891; APS_kinase.
CC DR Pfam; PF01583; APS_kinase; 1.
CC DR ProDom; PD002350; APS_kinase; 1.
CC DR TIGRFam; TIGR00455; apsk; 1.
CC DR ATP-binding; Chloroplast; Cysteine biosynthesis; Kinase;
KW phosphorylation; transferase; Transist peptide.
KW TRANSIT 1 ? Chloroplast (Potential).
FT CHAIN ? 312 Adenylyl-sulfate kinase.
FT NP BIND 142 149 ATP (Potential).
FT FT ACT_SITE 216 216 Phosphoserine intermediate (By
FT similarity).
SQ SEQUENCE 312 AA; 33656 MW; 16BBD11FB4B1FE27 CRC64;

Query Match 45.6%; Score 833; DB 1; Length 312;
Best Local Similarity 72.9%; Pred. No. 2.3e-53;
Matches 156; Conservative 24; Mismatches 34; Indels 0; Gaps 0;

Qy 130 PVTEKPMWNIKSTWILWHNCLIGSDQKLGQKGVVWITGLSGSGKSTLACALSRE 189
Db 99 PGKKILQTTTGVNSTWILWHKCAVEKSERQEPQQRGCVIWTGLSGSGKSTLACALSRG 158

Qy 190 LHCRGHLYTVLGDNLRHGLNRDLSFKAEDRAENIRRVGEVAKLFADAGVICIASLISPY 249
Db 159 LHAQKLYTILGDNVHGLNSDLSFKAEDRAENIRRIERGEVAKLPADAGVICIASLISPY 218

Qy 250 RRDRDRAALLPHSNFIEFIDLPLKICEARDPKGLYKLARTKIKGFTGIDDPYEPFIN 309
Db 219 RKPPDACSRLPEGDFIEVMDVPLKVEARDPKGLYKLARAGKIKGFTGIDDPYEPPLK 278

Qy 310 GEIVIKMKDECPSPKMAKQVLCYLENGYLQA 343
Db 279 SEIVLHQKLGWCDSPCDLADIVISYLENGYLKA 312

RESULT 4
KAP1_ARATH
ID KAP1_ARATH STANDARD; PRT; 276 AA.
AC Q43295;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Adenylyl-sulfate kinase 1, chloroplast precursor (EC 2.7.1.25) (APS
DE kinase) (Adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-
DE phosphosulfate 3'-phosphotransferase).

```



```

Oy 100 GARTCHRGIGRWVRRRRNGAAGEAP---HSPVKEKPVMSNIGKSTNHLWHNCLIGQS 156
Db 43 GSQTUSH-----NKNGSIPVKSINGHTGQKQP-LSTGVNSTNWKHCEKVKV 91

Oy 157 DRQKLLGKQGVVWITGLSGGSKSTLACALGRELHCRGHLTVYLDGDLNRLHGLNRLDSFK 216
Db 92 DRQRLDQKGVVWITGLSGGSKSTLACALNQLYKQKGLCYLDGDNVRHGLNRLDSFK 151

Oy 217 AEDRAENRRVGEVAKLFADAGVICIASLISPYRDRDRAALLPHSNFIEVFDLPLKI 276
Db 152 AEDRAENRRVGEVAKLFADAGVICIASLISPYRDRDRAALLPHSNFIEVFDLPLKI 211

Oy 277 CEARDPKGLYKLARTGKIGFTGIDDPYEPPIINGEIVIMKDEECPSPKAMAKOVLCYLE 336
Db 212 CEARDPKGLYKLARTGKIGFTGIDDPYEPPLNCE--ISLREGTSGTPIEMAEKVGVYLD 269

Oy 337 ENGYLQA 343
Db 270 NKGYLQA 276

RESULT 6
Q9SE92
ID Q9SE92 PRELIMINARY; PRT; 288 AA.
AC Q9SE92;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Adenosine-5'-phosphosulfate kinase (EC 2.7.1.25) (fragment).
GN Name=AK1;
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Bolchi A., Petrucci S., Ottonello S.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF178976; AAF17236.1; -.
DR HSSP; Q12657; IM70.
DR GO; GO:0004020; F:Adenylyl-sulfate kinase activity; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016772; F:transferase activity, transferring phosphor. .; IEA.
DR GO; GO:000103; P:sulfate assimilation; IEA.
DR InterPro; IPR002891; APS_kinase.
DR Pfam; PF01583; APS_kinase; 1.
DR ProDom; PD002350; APS_kinase; 1.
DR TIGRFAMs; TIGR00455; apsk; 1.
KW Kinase; Transferase.
FT NON_TER 1
SQ SEQUENCE 288 AA; 31612 MW; 291AE77AD184FF07 CRC64;

Query Match 44.5%; Score 812.5; DB 2; Length 288;
Best Local Similarity 53.9%; Pred. No. 6.9e-52;
Matches 172; Conservative 34; Mismatches 62; Indels 51; Gaps 7;

Oy 40 PPTLAVILVN-----PQAPVPLGLTSDAPLPAVLVHGTTPRSSHSSAGLASDSGR 94
Db 2 PVPVLLAVARASRAPEPEPVKGRVASAPPPPE-----TP-----SGWN 43

Oy 95 EGEGRGARTHCHRGIGRWVRRRRNG-----AAPGEAPHSPVKEK--PVMSNIGK 142
Db 44 QNEGD-----RRKRGTPVLECTGDSRVEEQFEHAGVEGKAWKMSSTVPK 89

Oy 143 STNHLWHNCLIGSRQKLLGKGVVWITGLSGGSKSTLACALGRELHCRGHLTVYLDG 202
Db 90 SSNIFWHDQKGVVWITGLSGGSKSTLACALNQLYKQKGLCYLDGDNVRHGLNRLDSFK 149

Oy 203 DNLHGLNRLDSFKAEADRAENRRVGEVAKLFADAGVICIASLISPYRDRDRAALLPH 262
Db 150 DNLHGLNKLQGLKAEADRAENRRV--VAKLFADAGLVCIASLISPHRRDRESCALLSD 207

```

```

Oy 263 SNFIEVFDLPLKICEARDPKGLYKLARTGKIGFTGIDDPYEPPIINGEIVIMKDEEC 322
Db 208 SSFIEVLNLSLEICEARDPKGLYKLARTGKIGFTGIDDPYEPPLNCEIIEIKVEDGVCP 267

Oy 323 SPKAMAKOVLCYLENGYL 341
Db 268 PPAEMAGQVVTYLEKGF 286

RESULT 7
KAP2 ARATH
ID KAP2 ARATH STANDARD; PRT; 293 AA.
AC Q49196;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Adenylyl-sulfate kinase 2, chloroplast precursor (EC 2.7.1.25) (APS kinase) (Adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-phosphotransferase).
GN Name=AKN2; OrderedLocusNames=At4G39940; ORFNames=T5J17.110;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Schiffrmann S., Schwenn J.-D.;
RT "Isolation of cDNA clones encoding adenosine-5'-phosphosulfate-kinase (EC2.7.1.25) from Catharanthus roseus and an isoform (akn2) from Arabidopsis.";
RL (er) Plant Gene Register PGR98-116.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Schiffrmann S., Schwenn J.-D.;
RT "Isolation of cDNA clones encoding adenosine-5'-phosphosulfate-kinase (EC2.7.1.25) from Catharanthus roseus and an isoform (akn2) from Arabidopsis.";
RL (er) Plant Gene Register PGR98-116.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Mayer K.F.X., Schueller C., Mambutt R., Murphy G., Volckaert G., Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N., Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M., Weichselgarthner M., de Simone V., Obermaier B., Mache R., Mueller M., Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidtkeini T., Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I., Vos P., Hohesl J., Zimmermann W., Wedler H., Ridley P., Langham S.-A., McCullagh B., Bilham L., Robben J., Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F., Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E., Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M., Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W., Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P., Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H., De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R., Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S., Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R., Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S., Borkova D., Bloeker H., Schaefer M., Grimm M., Loeuher T.-H., Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fartmann B., Grandjean K., Dauner D., Herzl A., Neumann S., Argirou A., Vitale D., Liguori R., Piravandi E., Massenot C., Quigley F., Clabaud G., Muendlein A., Felber R., Schnabl S., Hiller R., Schmidt W., Lechamy A., Aubourg S., RA Chedof F., Cooke R., Berger C., Monfort A., Casacuberta E., RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A., Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T., Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C., Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S., Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K., Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L., Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J., Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J., Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D., Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K., Kramer J., Fulton L., Mardis E., Dance M., Pepin K., Hillier L.W.,

```


RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antonio B., Zidan M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vill D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tull S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Marra M.A., Martienssen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
thaliana";
RL Nature 402:769-777(1999).
CC -!- FUNCTION: Catalyzes the synthesis of activated sulfate.
CC -!- CATALYTIC ACTIVITY: ATP + adenylyl sulfate = ADP + 3'-
CC phosphoadenylyl sulfate.
CC -!- PATHWAY: Sulfate activation; cysteine biosynthesis reductive
CC branch; second step.
CC -!- SUBCELLULAR LOCATION: Chloroplast (By similarity).
CC -!- SIMILARITY: Belongs to the APS kinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF043351; AAC39520.1; -;
CC EMBL; AL035708; CAB38907.1; -;
CC EMBL; AL161596; CAB80657.1; -;
CC PIR; T06100; T06100.
CC HSSP; Q12657; IM7G.
CC InterPro; IPR002891; APS_kinase.
CC Pfam; PF01583; APS_kinase; 1.
CC ProDom; PD002350; APS_kinase; 1.
CC TIGRfam; TIGR00455; apsk; 1.
CC
CC ATP-binding; Chloroplast; Cysteine biosynthesis; Kinase;
KW Phosphorylation; Transferrase; Transit peptide.
FT TRANSIT ?
FT CHAIN ? 293 Chloroplast (Potential).
FT NP BIND 142 149 ATP (Potential).
FT ACT_SITE 136 136 Phosphoserine intermediate (By
FT similarity).
SQ SEQUENCE 293 AA; 31977 MW; 71FAD6E9B026886 CRC64;
Query Match 42.6%; Score 777.5; DB 1; Length 293;
Best Local Similarity 71.8%; Pred. No. 2.7e-49;
Matches 145; Conservative 26; Mismatches 30; Indels 1; Gaps 1;
QY 142 KSTNLHWHNCLIGQSDRQKLLGQKGVVWITGLSGSGKSTLACALSRELHCRGHLTYVLD 201
DB 91 KAENIVHHESSICRCDRQLQKQKGVVWITGLSGSGKSTVACALSRELHCRGHLTYVLD 150
QY 202 GNLRLHGLNRLDSFKAEDRAENIRRVGEVAKLPADAGVICTASLISPYRRDRDRAALLP 261
DB 151 GDNVHGLNRLDTFKAERHTNIRIRIGEVAKLFDVGVGVTICIASLISPYRRDRDRAALLP 210
QY 262 HSNFTLVFIDLPKICEARDPKGLYKLTARKTGKGTGIDDPYEPPIGEIVIK-WKDEE 320
DB 211 DGDFFVEFMDVPLHVCESRDPKGLYKLTARKTGKGTGIDDPYEPPIGEIVIK-KHTGDDE 270
QY 321 CPSKMAKQVLCYLEENGYLQ 342
DB 271 SCSPRQMAENIISYLNQKYLE 292
RESULT 8
Q84JF0
ID Q84JF0 PRELIMINARY; PRT; 310 AA.
AC Q84JF0
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative adenylylsulfate kinase.

GN Name=At5g67520;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Huan V.W., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,
RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Huan V.W., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,
RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT0031977; AAO42019.1; -;
DR EMBL; BT005193; AAO50726.1; -;
DR HSSP; Q12657; IM7G.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016772; F:transferase activity, transferring phosphor. .; IEA.
DR GO; GO:000103; P:sulfate assimilation; IEA.
DR InterPro; IPR002891; APS_kinase.
DR Pfam; PF01583; APS_kinase; 1.
DR ProDom; PD002350; APS_kinase; 1.
DR TIGRfam; TIGR00455; apsk; 1.
KW Kinase.
SQ SEQUENCE 310 AA; 34064 MW; CIAC01A4242A0BDF CRC64;
Query Match 41.5%; Score 757.5; DB 2; Length 310;
Best Local Similarity 66.4%; Pred. No. 8.6e-48;
Matches 144; Conservative 29; Mismatches 29; Indels 15; Gaps 2;
QY 141 GKSTNLHWHNCLIGQSDRQKLLGQKGVVWITGLSGSGKSTLACALSRELHCRGHLTYVLD 200
DB 84 GKQNNVHWHDCPVTKSDRQELIKQKGVVWITGLSGSGKSLACALSRELHCRGHLTYVLD 143
QY 201 DGDNLHGLNRLDSFKAEDRAENIRRVGEVAKLPADAGVICTASLISPYRRDRDRAALLP 260
DB 144 DGDNVHGLNRLDSFKAEDRAENIRRVGEVAKLPADAGVICTASLISPYRRDRDRAALLP 203
QY 261 PHSNFTLVFIDLPKICEARDPKGLYKLTARKTGKGTGIDDPYEPPIGEIVIK----- 315
DB 204 PQGDFIEFMDVPLHVCESRDPKGLYKLTARKTGKGTGIDDPYEPPIGEIVIK-SDK 263
QY 316 -----MKDEECPSPKA-----MAKQVLCYLEENGYLQ 342
DB 264 GLSSSSSSSSSSSSSSSLCEMADIWVSYLQNGYLK 300
RESULT 9
Q9FJX1
ID Q9FJX1 PRELIMINARY; PRT; 290 AA.
AC Q9FJX1
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Adenylylsulfate kinase-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.

```
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Sequence features of the regions of 1,367,185 bp covered by 19
RT physically assigned pl and TAC clones.";
RL DNA Res. 5:203-216(1998).
DR EMBL; AB013390; BAB08460.1; -.
DR HSSP; Q12657; 1M7G.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016772; F:transferase activity, transferring phosphor. .; IEA.
DR GO; GO:0000103; P:sulfate assimilation; IEA.
DR GO; GO:0000103; P:sulfate assimilation; IEA.
DR InterPro; IPR002891; APS_kinase.
DR Pfam; PF01583; APS_kinase; 1.
DR ProDom; PD002350; APS_kinase; 1.
DR TIGRFAMs; TIGR00455; apsK; 1.
KW Kinase.
SQ SEQUENCE 290 AA; 32195 MW; BB7FB12990A6F6F8 CRC64;

Query Match 41.3%; Score 755; DB 2; Length 290;
Best Local Similarity 69.8%; Pred. No. 1.2e-47;
Matches 141; Conservative 28; Mismatches 33; Indels 0; Gaps 0;

QY 141 GKSTNLWHNCLIGOSDRQKLLGKGVVWITGLSGSKSTLACALSRELHCRGHLTYVL 200
DB 79 GKQKNIVVHDCPVTKSDRQELIKHGCVIWTGLSGSKSLACALSRELHNRGKLSYIL 138

QY 201 DGDNLRLHGLNRLSFKAEADRAENIRRVGEVAKLFADAGVICIASLISPYRRDRACRALL 260
DB 139 DGDNVRHGLNSDLSEADRAENIRRVGEVAKLFADSGIICIASLISPYRIERACRALL 198

QY 261 PHSNFIIEVDLPLKICEARDPKGLYKLRAGKIKGFTGIDDPYEPPIINGEIVIKMDEE 320
DB 199 PQGDFIEVFMVDVPLHVCEARDPKGLYKRRAGKIKGFTGVDVDPYEPALDCEVHIISNFS 258

QY 321 CPSPKAMAKVLCYLENGYLQ 342
DB 259 SSSLCEMADIVVSYLDQNGYLK 280

RESULT 10
Q8LF64 PRELIMINARY; PRT; 305 AA.
AC Q8LF64;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Adenylsulfate kinase-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22088475; PubMed=12093376;
RA Haas B.J., Voifovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY085031; RAM61589.1; -.
DR HSSP; Q12657; 1M7G.
DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016772; F:transferase activity, transferring phosphor. .; IEA.
DR InterPro; IPR002891; APS_kinase.
DR Pfam; PF01583; APS_kinase; 1.
DR ProDom; PD002350; APS_kinase; 1.
DR TIGRFAMs; TIGR00455; apsK; 1.
FT NON TER 152 152
SQ SEQUENCE 152 AA; 16974 MW; DBAC182B71D50DES CRC64;

Query Match 34.6%; Score 632; DB 2; Length 152;
Best Local Similarity 77.0%; Pred. No. 6.5e-39;
Matches 117; Conservative 20; Mismatches 15; Indels 0; Gaps 0;

QY 137 MSNIGKSTNLWHNCLIGOSDRQKLLGKGVVWITGLSGSKSTLACALSRELHCRGHL 196
DB 1 MSTVGNSTNIFWQESPFGKTRQKLLNQKGVVWITGLSGSKSTLACALSRELNRGKL 60
```

[illegible]

QY	316	MKDDECPSPKAMAKQVLCYLEENGYL	341
		: : : :	
DB	181	DTEKCD-----VDNIIIEYLKOKHIL	200

OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Cionidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP SEQUENCE FROM N.A.
RA Hotta K., Takahashi H., Asakura T., Saitoh B., Takatori N., Satou Y.,
RA Satoh N.;
RL Submitted (JAN-2000) to the ENML/GenBank/DBJ databases.

DR Pfam; PF01583; APS kinase; 1.
DR Pfam; PF01747; ATP-sulfurylase; 1.
DR TIGRFAMs; TIGR00455; apsk; 1.
KW Kinase.
SQ SEQUENCE 618 AA; 69559 MW; 75223CCB022BC297 CRC64;

Matches 114; Conservative 30; Mismatches 51; Indels 14; Gaps 5

Db 2 ASATANHGDIDIQEPPM-----SSNIYVQQHVSRDERGQVMGKCAFGRCTVWFGLSGA 56

Qy 178 GKSTLACALSRHLHCRGHLTVYLDGDNLRHGLNRDLSPKADRDRAENIRRVGEVAKLFADA 237

Db 57 GKSTLAKMLREYLCCKGIPAYSLDGDNI RHGLNRDLGAPADREENIRRI GEVAKLFADA 116

```

db 117 GWCLVSFISPYRKOROSARRVHNKSNLPFTEVFVDTPQVCEGRDVKGLYKARQGIK 176
Qy 296 GFTGIDDPPEPPINGEIVIKMD---EEC 321
Db 177 GFTGISPYEPPENAEVTVTADMPVEDC 205

```

```
Db      177 GFtGIDSPYEPPENAEVVTKTADMPVEDC 205
                                     :
RESULT 14
CYC2 BACSU
ID _CYC2 BACSU          STANDARD;          PRT;   199 AA.
AC O06735;
```

DR	Subtilist; BG13105; yisz.
DR	HAMAP; MF_00065; -; 1.
DR	InterPro; IPRO02891; APS kinase.
DR	Pfam; PF01583; APS kinase; 1.
DR	ProDom; PD002350; APS kinase; 1.
DR	TIGRFAMS; TIGR00455; apsk; 1.
DR	ATP-binding; Complete proteome; Cysteine biosynthesis; Kinase;
KW	Phosphorylation; Transferase.
FT	NP_BIND 34 41 ATP (By similarity).
FT	ACT_SITE 108 108 Phosphoserine intermediate (By similarity)
FT	SEQUENCE 199 AA; 22304 MW; 90EC37FE4B02A123 CRC64;
SQ	Query Match 30.4%; Score 555; DB 1; Length 199;
	Best Local Similarity 58.5%; Pred. No. 4.3e-33;
	Matches 100; Conservative 27; Mismatches 44; Indels 0; Gaps 0;
Qy	145 NIIWHNCLIGSQDRQKLLGQGVVITGLSGSKSTLACALSRELHCRGHLYVLGDGN 204
Db	6 NIITHPAALISKSDRSINGHKSCVLNFTGLSGSKSVLANAVDEKLYRKGIQSIVLDGDN 65
Qy	205 LRHGINDLSFKAEDRAENIRRVGEYAKLFADAGVICIASLSPYRRDRDACALLPHSN 264
Db	66 IREGLNKDLGFQDGRDENIRRGVEYAKLVDSGGOMILTAFFSPPREDRMVRALPFKGE 125
Qy	265 FIEVFIDLPKICEARDPKGLYKLARTGKIKGFTGIDDVPPEPPINGSIIVK 315
Db	126 FFBIYVKPLHVCEQRDPKGLYKARNGEIKHFTGIDSPEAPLSPDFIIE 176
RESULT 15	
Q73WR1	PRELIMINARY; PRT; 230 AA.
ID Q73WR1	
AC Q73WR1	
DT 05-JUL-2004	(TEMBLrel. 27, Created)
DC 05-JUL-2004	(TEMBLrel. 27, Last sequence update)
DT 05-JUL-2004	(TEMBLrel. 27, Last annotation update)
DE Hypothetical protein.	
GN OrderedLocusNames=MAP2599c;	
OS Mycobacterium paratuberculosis.	
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;	
CC Corynebacterinae; Mycobacteriaceae; Mycobacterium.	
OX NCBI_TaxId=1770;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=k10;	
RA Li L., Barnantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;	
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.	
EMBL; AE017236; AAS04916.1; -.	
GO; GO:0005524; F:ATP binding; IEA.	
DR GO; GO:0016301; F:kinase activity; IEA.	
DR GO; GO:0016772; F:transferase activity, transferring phosphor. . ; IEA.	
DR GO; GO:0000103; P:sulfate assimilation; IEA.	
DR InterPro; IPRO02891; APS kinase.	
DR Pfam; PF01583; APS kinase; 1.	
DR ProDom; PD002350; APS kinase; 1.	
DR TIGRFAMS; TIGR00455; apsk; 1.	
KW Complete proteome; Hypothetical protein.	
SQ SEQUENCE 230 AA; 25507 MW; 3FCDA4A2FC214DD8 CRC64;	
SQ	
Query Match 30.4%; Score 555; DB 2; Length 230;	
Best Local Similarity 50.5%; Pred. No. 5.1e-33;	
Matches 109; Conservative 34; Mismatches 53; Indels 20; Gaps 2;	
Qy	143 STNILWHNCLIGSQDRQKLLGQGVVITGLSGSKSTLACALSRELHCRGHLYVLVDG 202
Db	2 SNNITWHEIKISGREQLNGHKCVIWFITGLSGSKSTVANVVEQKLYERGIRSYLLDG 61
Qy	203 DNLRHGLNR-----DLSFKADRAENIRRVGEYAKLFADAGVICIASL 245
Db	62 DNVRYGLNAGPDLLERHGPFAQRFLGFGSAODREENIRRI GAVAKLFCFEGAIALTAF 121
Qy	246 ISPYRRDRACRALPHSNFIEVFIDLPIKCIPARDPKGLYKLARTGKIKGFTGIDDPYE 305

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 9, 2005, 13:07:17 ; Search time 665 Seconds
(without alignments)
3387.314 Million cell updates/sec

Title: US-10-829-432-4

Perfect score: 1826

Sequence: 1 RPFPHINQTEPLVTHQTPP.....PRMAKQVLCYLENGYLQA 343

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 7351250 seqs, 3283620254 residues

Total number of hits satisfying chosen parameters: 14702500

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10829432/runat_08092005_132224_25132/app_query.fasta.1.519
-DB=Published Applications NA -QMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10829432.cgn_1.1.723 @runat_08092005_132224_25132
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORE=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:
22: /cgn2_6/ptodata/1/pubpna/US10J_NEW_PUB.seq:
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1826	100.0	1217	19	US-10-829-432-3	Sequence 3, Appli
2	1660.5	90.9	1291	20	US-10-425-115-116953	Sequence 116953,
3	1415	77.5	1335	18	US-10-425-114-941	Sequence 941, App
4	1250.5	68.5	1498	20	US-10-425-115-116950	Sequence 116950,
5	1008.5	55.2	1319	20	US-10-739-930-5324	Sequence 5324, Ap
6	944	51.7	945	19	US-10-437-963-51779	Sequence 51779, A
7	936	51.3	945	19	US-10-767-701-9302	Sequence 9302, Ap
8	905.5	49.6	1028	18	US-10-425-114-16540	Sequence 16540, A
9	896.5	49.1	2611	20	US-10-425-115-62889	Sequence 62889, A
10	892.5	48.9	1180	18	US-10-425-114-18247	Sequence 18247, A
11	887	48.6	890	19	US-10-829-432-1	Sequence 1, Appli
12	876	48.0	928	19	US-10-829-432-9	Sequence 9, Appli
13	875.5	47.9	1312	18	US-10-425-114-1150	Sequence 1150, Ap
14	854.5	46.8	1764	20	US-10-425-115-135127	Sequence 135127,
15	854	46.8	1116	18	US-10-425-114-9652	Sequence 9652, Ap
16	854	46.8	1256	18	US-10-424-599-62905	Sequence 62905, A
17	853	46.7	981	18	US-10-424-599-119620	Sequence 119620,
18	852	46.7	936	19	US-10-829-432-7	Sequence 7, Appli
19	846.5	46.4	1302	18	US-10-425-114-27400	Sequence 27400, A
20	816	44.7	831	9	US-09-938-842A-710	Sequence 710, App
21	816	44.7	831	11	US-09-938-842A-710	Sequence 710, App
22	788	43.2	687	17	US-10-260-238-869	Sequence 869, App
23	783.5	42.9	1821	20	US-10-425-115-35035	Sequence 35035, A
24	778.5	42.6	920	16	US-10-425-115-35036	Sequence 35036, A
25	770.5	42.2	548	19	US-10-437-963-72904	Sequence 65, Appl
26	690	37.8	1132	18	US-10-424-599-119621	Sequence 72904, A
c	640	35.0	378	17	US-10-260-238-5283	Sequence 119621,
28	615.5	33.7	635	20	US-10-425-115-35035	Sequence 5283, Ap
29	571	31.3	603	17	US-10-282-122A-15460	Sequence 15460, A
c	561	30.7	595	17	US-10-260-238-4843	Sequence 4843, Ap
31	555	30.4	600	17	US-10-369-493-41371	Sequence 41371, A
32	534.5	29.3	609	17	US-10-369-493-40905	Sequence 40905, A
33	530	29.0	1592	19	US-10-437-963-45866	Sequence 45866, A
34	521.5	28.6	597	17	US-10-282-122A-34804	Sequence 34804, A
35	521.5	28.6	615	22	US-10-724-972A-1295	Sequence 1295, Ap
36	519.5	28.5	588	17	US-10-369-493-40371	Sequence 40371, A
37	516	28.3	1863	17	US-10-369-493-34432	Sequence 34432, A
38	514.5	28.2	591	17	US-10-282-122A-10327	Sequence 10327, A
39	514.5	28.2	612	17	US-10-282-122A-32765	Sequence 32765, A
40	510.5	28.0	1842	17	US-10-369-493-39203	Sequence 39203, A
41	510.5	28.0	1869	17	US-10-369-493-39946	Sequence 39946, A
42	510.5	28.0	1890	17	US-10-369-493-39575	Sequence 39575, A
43	510	27.9	642	17	US-10-282-122A-41542	Sequence 41542, A
44	505.5	27.7	597	17	US-10-282-122A-35733	Sequence 35733, A
45	503	27.5	603	17	US-10-282-122A-22900	Sequence 22900, A

ALIGNMENTS

RESULT 1
US-10-829-432-3
; Sequence 3, Application US/10829432
; Publication No. US20040177401A1
; GENERAL INFORMATION:
; APPLICANT: Falco, Saverio
; APPLICANT: Allen, Stephen
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-B
; CURRENT APPLICATION NUMBER: US/10/829,432
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: US/09/720,384A
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/092,833
; PRIOR FILING DATE: 1998-07-14

```
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 1217
; TYPE: DNA
; ORGANISM: Zea mays
US-10-829-432-3
Alignment Scores:
Pred. No.: 4 44e-168 Length: 1217
Score: 1826.00 Matches: 343
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

US-10-829-432-4 (1-343) x US-10-829-432-3 (1-1217)
QY 1 ArgProPheHisPheIleAsnGlnThrGluProLeuValThrHisThrGlnGlnProPro 20
DB 2 CGTCGGTTTCATTTTCATCAATCAACAGAAACCTCTGGTCACACACAGCAGCAACACCG 61
QY 21 SerProAlaProGlyProHisSerGlnGlnArgGlnGlnGlnGlnGlnGlnGlnGln 40
DB 62 AGCCCGAGCGCCCGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 121
QY 41 ThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaProValLeuProGly 60
DB 122 AGCCGACGCTCGCGCTCATCTCTGTAATCCACAGCGCGCCCTCCGCTCCACAGC 181
QY 61 LeuThrProSerAspAlaProLeuProAlaLeuValIleHisGlyLeuThrProArgSer 80
DB 182 CTCACCCCTAGCGATGCGCCACTCCCGCGGCTCGTGATCATGGCTCACTCCCGCTCC 241
QY 81 SerHisSerSerAlaGlyLeuAlaSerAspSerGlyArgArgGluGlyGluGlyArgGly 100
DB 242 TCACACTCTTCGCGGGTCTCCCACTGATAGTGGCGCGCGCGAGGGAGGGCGCGGT 301
QY 101 AlaArgThrHisCysHisArgGlyIleGlyArgTTPValArgArgArgArgAsnGly 120
DB 302 GCGGTACGCACTGCCACCGCGCATTTGGCGGTGGTGGCGGGCGGGCGGGAGATGA 361
QY 121 AlaAlaProGlyGluAlaProHisSerProValLysGluLysGluProValMetSerAsnIle 140
DB 362 GCAGCGCCCGGGAGGCCCGCACAGCCCGAGTGAAGGAGAGCCCTGTAATGTCGAACATT 421
QY 141 GlyLysSerThrAsnIleLeuThrHisAsnCysLeuIleGlyGlnSerAspArgGlnLys 160
DB 422 GGGAAATCGCACTAATATTTATGGCACAATTCCTTGGTGGCAATCTGTATAGACAGAAA 481
QY 161 LeuLeuGlyGlnLysGlyCysValValTTPileThrGlyLeuSerGlySerGlyLysSer 180
DB 482 TTGCTGGGACAAAGAGCTGTGTGTATGATATACAGGACTCAGTGGTTACGGGAAAGT 541
QY 181 ThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeu 200
DB 542 ACTCTTGATGTGCACTGAGTGTGAGTTGCATTCGAGAGGCCACCTCAGCTATGTACTT 601
QY 201 AspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheIysAlaGluAspArg 220
DB 602 GATGTGACAACTTCAGACATGGCTTAAATAGAGATTTTAAGCTTTTAAGGACAGAACCGT 661
QY 221 AlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIle 240
DB 662 GCAGAAAATATACGAAGAGTTGGTGAAGTGGCAAGCTTTTGTGTATGCTGTGTAT 721
QY 241 CysIleAlaSerIleLeuSerProTyrArgArgAspArgAspAlaCysArgAlaLeuLeu 260
DB 722 TGCAATGTGTAGCTTGATATCTCCATACAGAGAGATCGTGATGATCGCGTCTACTT 781
QY 261 ProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArg 280
DB 782 CCACATCTTAACCTTTATGAAGTATTTATGTATTTTGGCCCTTAAATAATTTGTGAAGCTCGT 841

; US-10-829-432-4 (1-343) x US-10-425-115-116953 (1-1291)
QY 16 ThrGlnGlnProProSerProAlaProGlyProAlaSerGlnGlyGlnArgGlnGlyAsn 35
DB 1 AGCGAGCAACACACCGAGCGCGCCCGCCGCCACCCAC-----GGCAAC 44
QY 36 ThrLeuLeuSerProThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaPro 55
DB 45 ACCCTCCTCAGCCGAGCGCGCGCTCGCGCTCATCTCGTAAATCCACAGCGCGCGCT 104
QY 56 ProValLeuProGlyLeuThrProSerAspAlaProLeuProAlaLeuValIleHisGly 75
DB 105 CCGGCTCCCGAGGTCTCACCCTTAGCATGCGCCACTCCCGCGCTCGTATCATGCGC 164
QY 76 LeuThrProArgSerHisSerAlaGlyLeuAlaSerAspSerGlyArgArgGlu 95
DB 165 CTCACCTCCCGTCTCTCACACTCTTCGCGGGCGCTCCGCGAGATAGTGGCGCGCGAG 224
QY 96 GlyGluGlyArgGlyAlaArgThrHisCysHisArgGlyIleGlyArgTyrValArgArg 115
DB 225 GGGAGGGCGCGGTGCGGTACGCACTGCCACCGCGCATTTGGCGGTGGGTGGCGCGG 284
QY 116 ArgArgArgAsnGlyAlaAlaPro-GlyGluAlaProHisSerProValLysGluLysPr 135
DB 285 CGCGCGGAATGGAGCGCGCCACGGGAGGGCCCGCCAGCCCGCCAGCCAGGAGAGGCC 344

RESULT 2
US-10-425-115-116953
; Sequence 116953, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kowalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 116953
; LENGTH: 1291
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_38151C.1
US-10-425-115-116953
Alignment Scores:
Pred. No.: 6 47e-152 Length: 1291
Score: 1660.50 Matches: 319
Percent Similarity: 97.57% Conservative: 2
Best Local Similarity: 96.96% Mismatches: 2
Query Match: 90.94% Indels: 7
DB: 20 Gaps: 1

US-10-829-432-4 (1-343) x US-10-425-115-116953 (1-1291)
QY 16 ThrGlnGlnProProSerProAlaProGlyProAlaSerGlnGlyGlnArgGlnGlyAsn 35
DB 1 AGCGAGCAACACACCGAGCGCGCCCGCCGCCACCCAC-----GGCAAC 44
QY 36 ThrLeuLeuSerProThrProThrLeuAlaValIleLeuValAsnProGlnArgAlaPro 55
DB 45 ACCCTCCTCAGCCGAGCGCGCGCTCGCGCTCATCTCGTAAATCCACAGCGCGCGCT 104
QY 56 ProValLeuProGlyLeuThrProSerAspAlaProLeuProAlaLeuValIleHisGly 75
DB 105 CCGGCTCCCGAGGTCTCACCCTTAGCATGCGCCACTCCCGCGCTCGTATCATGCGC 164
QY 76 LeuThrProArgSerHisSerAlaGlyLeuAlaSerAspSerGlyArgArgGlu 95
DB 165 CTCACCTCCCGTCTCTCACACTCTTCGCGGGCGCTCCGCGAGATAGTGGCGCGCGAG 224
QY 96 GlyGluGlyArgGlyAlaArgThrHisCysHisArgGlyIleGlyArgTyrValArgArg 115
DB 225 GGGAGGGCGCGGTGCGGTACGCACTGCCACCGCGCATTTGGCGGTGGGTGGCGCGG 284
QY 116 ArgArgArgAsnGlyAlaAlaPro-GlyGluAlaProHisSerProValLysGluLysPr 135
DB 285 CGCGCGGAATGGAGCGCGCCACGGGAGGGCCCGCCAGCCCGCCAGCCAGGAGAGGCC 344
```



```
QY 135 oValMetSerAsnIleGlyLeuSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyG1 155
Db 345 TGTAAATGTCACCAATTCGGGAATCGACTAATATTTATGACCAATTCGTCATGGACA 404
QY 155 nSerAspArgGlnLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSe 175
Db 405 ATCTGATAGACAGAAATTCCTGGGCAAAAAGGCTGTGTCGATAGGATAACAGGACTCAG 464
QY 175 rGlySerGlyLeuSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHi 195
Db 465 TGGTTCAGGAAAAGTACTCTTGCATGTGCACTGAGTCGTGAGTTGCATTCAGAGGCCA 524
QY 195 sLeuThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPh 215
Db 525 CCTCAGTATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 584
QY 215 eLysAlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAl 235
Db 585 TAAGGAGAGAACCGCTGCAGAAAATATACGAAAGTTGTTGAAGTGGCAAGCTTTTTC 644
QY 235 aAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgArgAspAl 255
Db 645 TGATGCTGTGTGTCATGATGTTGCTAGCTTGTCTTCTCTCTCTCTCTCTCTCTCTCT 704
QY 255 cCysArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuLeuLy 275
Db 705 ATGCGTGTCTTACTTCCACATTCCTAACTTTATGAAGTATTTATGATTCGCCCTAAA 764
QY 275 sLysCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLy 295
Db 765 AATTTGTGAAGCTCGTGATCTCTAAAGGCTATACAAAGCTTGCAAGCTTGCAAGAAAT 824
QY 295 sGlyPheThrGlyIleAspAspProTyrGluProProIleAsnGlyGluIleValIleLy 315
Db 825 AGTTTCTGCAATTTGATGATCCATACGAAACCCCAATTAATGATGATGATGATGATGAT 884
QY 315 sMetLysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrIle 335
Db 885 GATGAAGATGGGAATGCCCTTCCACCAAGCAATGGCCAGCAAGTTCATGCTACTCT 944
QY 335 uGluGluAsnGlyTyrLeuGlnAla 343
Db 945 TGAAGAGAACGGATATTTGCAAGCT 969
```

RESULT 3

```
US-10-425-114-941
; Sequence 941, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 941
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700097023_FLI
US-10-425-114-941
```

Alignment Scores:

```
Pred. No.: 5,76e-128 Length: 1335
Score: 1415.00 Matches: 278
```

```
Percent Similarity: 90.00% Conservative: 1
Best Local Similarity: 89.68% Mismatches: 3
Query Match: 77.49% Indels: 28
DB: 18 Gaps: 1
US-10-829-432-4 (1-343) x US-10-425-114-941 (1-1335)
QY 62 ThrProSerAspAlaProLeuProAlaLeuValIleHisGlyLeuThrProArgSerSer 81
Db 2 ACCCTAGCGATGGCCACTCCCGCGCTCGTGATCCATGGCTCACTCCCGGTTCTCTCA 61
QY 82 HisSerSerAlaGlyLeuAlaSerAspSerGlyArgArgGluGlyGlyArgGlyAla 101
Db 62 CACTCTTCGGCGGCTCGCCAGCGATAGTGGCGCCGAGGGGGAGGCGCGCTGG 121
QY 102 ArgThrHisCysHisArgGlyIleGlyValArgTrpValArgArgArgAsnGlyAla 121
Db 122 CGTACGCACTGCCACCGCGCATTTGGCGGTAGGTGGCGCGCGCGCGGGAATGAGCA 181
QY 122 AlaPro-GlyGluAlaProHisSerProValLysGlyLysProValMetSerAsnIleG1 141
Db 182 GCGCCACGCGGAGGCGCCGACAGCCAGTGAAGGAGAAGCCTGTAATGTCAACATTGG 241
QY 141 YlySerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysIle 161
Db 242 GAAATCGACTAATATTTTATGGCAATTTGCTTGTATGGACAATCTGATAGACAGAAAT 301
QY 161 uLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerTh 181
Db 302 GCTGGGCAAAAAGGCTGTGCTGATGATTAACAGACTCAGTGTGTCAGGGAAGAAGTAC 361
QY 181 rLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAs 201
Db 362 TCTTGATGTGCACTGAGTGTGATGTCATTGTCAGAGGCCACCTCAGTATGACTTGA 421
QY 201 pGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAl 221
Db 422 TGGTGACAACTCAGACATGCGCTAAATAGAGATTTAAGCTTTAAGGCGAGAACCGTGC 481
QY 221 aGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCy 241
Db 482 AGAAATATACGAGAGTGTGTAAGTGGCAAGCTTTTCTGATGCTGGTGTCTATATG 541
QY 241 sIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeuLeuPr 261
Db 542 CATTCGTAGCTTGATATCTCCATACAGAGAGATGCTGATGATGCCGCTCTACTTCC 601
QY 261 oHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAs 281
Db 602 ACATTTCTAACTTTATTGAAGTATTTATTGATTTGCCCTTAAATAATTTGTGAAGCTCG 661
QY 281 pProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyLeas 301
Db 662 TCCTAAAGGCGCTGTACAAGCTTGCAAGCTTGCAAGGAAGATTAAGGTTTTCACCTGGA 721
QY 301 pAspProTyrGluProProIleAsnGlyGlu----- 311
Db 722 TGATCCATACGAAACCACTAATTAATGTTGAGATAAGGCAATAGTTATATACATCTGG 781
QY 312 -----lleValIleL 315
Db 782 TGTTTTCAAGGATTAACTTTGAACTCAGAACGCTCTTCTGTCAGTGTTAAGATAGTAATTA 841
QY 315 yMetLysAspGluCysProSerProLysAlaMetAlaLysGlnValLeuCystrL 335
Db 842 CGATGAAGATGGGGAATGCCCTTCCACCAAGCAATGGCCAAAGCAATGGCCAAAGTTC 901
QY 335 euGluGluAsnGlyTyrLeuGlnAla 343
Db 902 TTGAAGAGAACGGATATTTTGAAGCT 927
RESULT 4
US-10-425-115-116950
```

```
; Sequence 116950, Application US/10425115
; Publication NO. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 116950
; LENGTH: 1498
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_38149C.1
US-10-425-115-116950

Alignment Scores:
Pred. No.: 7,18e-112 Length: 1498
Score: 1250.50 Matches: 263
Percent Similarity: 78.70% Conservativeness: 3
Best Local Similarity: 77.81% Mismatches: 11
Query Match: 68.48% Indels: 64
DB: 20 Gaps: 3

US-10-829-432-4 (1-343) x US-10-425-115-116950 (1-1498)
QY 9 ThrGluProLeuValThrHisThrGlnGlnProProSerProAlaProGlyProAlaSer 28
DB 22 ACAGAACCTCTGTCA--CACAGCAGCAAGCAGCCAGCCAGCCGCCGCCAGCCAC- 78
QY 29 GlnGlyGlnArgGlnGlyAsnThrLeuLeuSerProThrProThrLeuAlaValleLeu 48
DB 79 -----GGCAACACCTCTCTCAGCCGAGCCGAGCGCTCGCGCTCATCCTC 123
QY 49 ValAsnProGlnArgAlaProValLeu-----ProGlyLeuThrProSerAspAla 66
DB 124 GTAAGTCCACAGCGCGCGCTCCCGTCTCCAGTCCCGAGGTCCTCACCCCTAGCGATGCG 183
QY 67 ProLeuProAlaLeuValleHisGlyLeuThrProArgSerSerHisSerSerAlaGly 86
DB 184 CCACCTCCCGCGCTCGTGATCATGGCTCATCTCCCGTCTCTCACACTTCCGCGGCG 243
QY 87 LeuAlaSerAspSerGlyArgArgGluGlyGluGlyArgGlyAlaArgThrHisCysHis 106
DB 244 CTCGCCAGCGATAGTGGCGCGCGCGAGGGGAGGGCGCGGTGCGGTACGACCTGCCAC 303
QY 107 ArgGlyVlleGlyArgTrpValArgArgArgArgAsnGlyAlaAlaPro-GlyGluAl 126
DB 304 CCGCGCATTTGGCGGTAGTGGCGCGCGCGCGCGGGAATGGAGCAGCGCCACGGGAGGC 363
QY 126 aProHisSerProValLeuGlyLeuProValMetSerAsnIleGlyLeuSerThrAsnI 146
DB 364 CCGCACAGCCAGTGAAGAGAGAGCGCTGTAATGTCAAACATTGGGAAATCGACTAATAT 423
QY 146 eLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuLeuGlyGlnLysGl 166
DB 424 TTTATGGCAAAATTGCTTATGGCAATCTGATAGACAGAAATTCGTGGGCAAAAAGG 483
QY 166 vCysValValTrpIleThrGlyLeuSerGlySerGlyLysSerThrLeuAlaCysAlaLe 186
DB 484 CTGTGTCGTATGATTAACAGGACTCAGTGTTCAGGAAAAGTACTCTTTCATGTGCACT 543
QY 186 uSerArgGluLeuHisCysArgGlyHisLeuThrTrpValLeuAspGlyAspAsnLeuAr 206
DB 544 GAGTCGTGAGTTGCAATTGTCAGAGGCCACCTCAGTATGTACTTGTATGGTGAACCTCAG 603
QY 206 gHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIleArgNr 226
DB -----
; Sequence 5324, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 5324
; LENGTH: 1319
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-CLUSTER7564_1
US-10-739-930-5324

Alignment Scores:
Pred. No.: 2,4e-88 Length: 1319
Score: 1008.50 Matches: 203
Percent Similarity: 80.80% Conservativeness: 20
Best Local Similarity: 73.55% Mismatches: 40
Query Match: 55.23% Indels: 13
DB: 20 Gaps: 4

US-10-829-432-4 (1-343) x US-10-739-930-5324 (1-1319)
QY 73 lIleHisGlyLeuThrProArgSerSerHisSerSerAlaGlyLeuAlaSerAspSerGly 92
DB 4 ATTCAC---GTCCAGCGCGCGCATTTTATTATTAGTCTCGCAGCAGCAGTGGCTGTGTA 60
QY 93 ArgArgGluGlyGluGlyArgGlyAlaArgThrHisCysHisArgGlylleGlyArgTrp 112
DB 61 CGAGTCGAGCAGCTCACGCGCGCGCGCGCTG---TGC-----TSG 99
QY 113 ValArgArgArgArgAsnGlyAlaAlaProGlyGluAlaProHisSerPro----- 130
DB 100 GAGTCTTCCTCCCTCCACTCCCGTCCAGCTCCCGGGCGCTCCCGCCGCCCCCACTC 159
QY 131 -----ValLysGluLysProValMetSerAsnIleGlyLysSerThrAsnIleLeu 147
DB 160 CGCGCGCGAGCTCGAGAGAGCGCTGTCATGTCAAACTTGGGAAATCACTAATATTCTA 219
QY 148 TrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuLeuGlyGlnLysGlyCys 167
DB -----
```

Db 220 TGGCATGCTGCCCAATTCGACAACTGACGACAGAAATTCGTGGCGCAAAAGGATGT 279
Qy 168 ValValTTPleThrGlyLeuSerGlySerGlySerThrlleuAlaCysAlaLeuSer 187
Db 280 GTATATGATACAGGATCTACGCGGTTTCAGGAAAGTACCGTTCCTGTGCTAGT 339
Qy 188 ArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAspGlyAspAsnLeuArgHis 207
Db 340 CGGGAATTACCTACAGAGGCCACACACGATATGCTCTTGATGGTGACAACTCAGACAT 399
Qy 208 GlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnLeuArgVal 227
Db 400 GGCCTCAATCGAGATCTAAGCTTCAAGGAGAGACCGTACAGAAAATATACGAAAGTT 459
Qy 228 GlyGluValAlaLysLeuPheAlaAspAlaGlyValLysCysLeuLeuSerLeuLeuSer 247
Db 460 GGTGAAGTGGCAAGCTTTTTCAGATGCTGTGTACCATATGATGCTAGTTGATATCT 519
Qy 248 ProTyrArgArgAspAlaCysArgAlaLeuLeuProHisSerAsnPheLeuGlu 267
Db 520 CCATACAGGAGACCGTATGATGTCGCGAGCTTACTTCCAGATCTAGATTTATTGAA 579
Qy 268 ValPheLeuAspLeuProLeuLysLysCysGluAlaArgAspProLysGlyLeuTyrLys 287
Db 580 GTATTTATGATTTGCCACTAGAAATTTATGTAAGCTCGTGATCTCTAAAGGGTTATACAAG 639
Qy 288 LeuAlaArgThrGlyLysLysGlyPheThrGlyLysAspAspProTyrGluProPro 307
Db 640 CTTGACGACAGGAAAGATTAAGGGTTTCACCGGAGTTGATGATCCATACGAATCACCA 699
Qy 308 IleAsnGlyGluLeuValLysMetLysAspGluGluCysProSerProLysAlaMet 327
Db 700 GTGATAGTGAGATAGTATTAGATGGAAGTGGGGAATGCCCTTCCAGGAAGCAATG 759
Qy 328 AlaLysGlnValLeuCysTyrLeuGluGluAsnGlyTyrLeuGlnAla 343
Db 760 GCCAGCAAGTTCTATCTCTACCTTGAGAAAGATGGATATTTGCAGGCT 807

RESULT 6

US-10-437-963-51779

; Sequence 51779, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 51779

; LENGTH: 915

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_5413C.1

US-10-437-963-51779

Alignment Scores:

Pred. No.:	3e-82	Length:	915
Score:	944.00	Matches:	194
Percent Similarity:	74.01%	Conservatives:	31
Best Local Similarity:	63.82%	Mismatches:	59
Query Match:	51.70%	Indels:	20
DB:	19	Gaps:	5

RESULT 7

US-10-767-701-9302

; Sequence 9302, Application US/10767701

; Publication No. US20040172684A1

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

US-10-829-432-4 (1-343) x US-10-437-963-51779 (1-915)

Qy 58 LeuProGlyLeuThrProSerAspAlaProLeuProAlaLeuValLeuHis----- 74
Db 5 TTGCGAGGATCTCCCATCCCGTCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 64
Qy 75 -----GlyLeuThrPro-ArgSerSerHisSerSerAlaGlyLeuAlaSer----- 89
Db 65 CTGCGCATGGCGCCTCTCCCGGAGGGGTACTGAGGCGCCACCGTGGGAGGAGTAGTGGCGG 124
Qy 90 -AspSerGlyArgArgGluGlyGlyArgGlyAlaArg-ThrHisCysHisArgGlyI 109
Db 125 GCGTGAAGCTGCTCGACGCGCGGCAATGAGCGTCTGACGCGCGCGCGCGCGCGCGGA- 183
Qy 109 leGlyArgTyrValArgArgArgArgAsnGlyAlaAla---ProGlyGluAlaProH 128
Db 184 --AGCAGCTGGACAGGAAGGCTGCGCGTCCGAGCGCGCGCGCGCGCGCGCGCG 241
Qy 128 isSer-----ProValLysGluLysProValMetSerAsnI 140
Db 242 TCGACGGGAAGGACGACGACGATGTGTTCAGATTTTAAACAGCTCATGACCTCACTG 301
Qy 140 leGlyLysSerThrAsnLeuLeuTrpHisAsnCysLeuLeuGlyGlnSerAspArgGlnL 160
Db 302 TCGGTAAATCAACAAACATCCGATGGCATGACTGCCAGTAAACCCAGCTTGACAGACAA 361
Qy 160 ysLeuLeuGlyGlnLysGlyCysValValTyrIleThrGlyLeuSerGlySerGlyLys 180
Db 362 AGTTCTCAACCCAGAAAGGGCTGTGTGTGGATCCTGCGCTAAAGTGGTTTCAGGAAAA 421
Qy 180 erThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrVal 200
Db 422 GCACCTTGCATGTGCATGAGCGCTGAGCTGCACTCAAGAGGGATCTGACCTATGTTTC 481
Qy 200 euAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 220
Db 482 TTGACGGCGACAATCTCCGCGCACGCGCTGAAACAAAGATCTCAGCTTCAAGGCCAAGATC 541
Qy 220 rgAlaGluAsnLeuArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 240
Db 542 GTGCGCAAAATATACGCGAGAGTTGAGAAAGTGGCAAAAGCTGTTCGACATGCTGGATTGA 601
Qy 240 leCysIleAlaSerLeuLeuSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 260
Db 602 TCTGCATTTACTAGTTTGTATATACCTATAGAGTATCGAAGCGCTGCGCGCAAAATCA 661
Qy 260 euProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 280
Db 662 TACCAATTCCTCGTTTTCATGAGGTGTTCGTAATGTCCCACTTGAAGTATGTGAAGAAA 721
Qy 280 rgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyI 300
Db 722 GGGATCCAAAAGGCGCTGACAAAGCTGCTGCGCGCGCAAAATCAAGGCTTTACGGGAA 781
Qy 300 leAspAspProTyrGluProLysLeuValLeuValLysMetLysAspGluG 320
Db 782 TAGATGATCCTTATGAAAACACCTTCAGATTGTGAGATTGTATACGATGCAAGTTGGGG 841
Qy 320 luCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluAsnGlyT 340
Db 842 ACTGCCCTTACCTAAATCAATGGCTGATCAAGTAGTGTCTATCTTGAAGCCATGGAT 901
Qy 340 yrLeuGln 342
Db 902 TCCTTCAG 909

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5335)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 9302
; LENGTH: 945
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS114933_1
US-10-767-701-9302

Alignment Scores:
Pred. No.: 1.88e-81 Length: 945
Score: 936.00 Matches: 178
Percent Similarity: 96.81% Conservative: 4
Best Local Similarity: 94.68% Mismatches: 6
Query Match: 51.28% Indels: 0
DB: 19 Gaps: 0

US-10-829-432-4 (1-343) x US-10-767-701-9302 (1-945)

```
QY 156 SerAspArgGlnLysLeuGlyGlnLysGlyCysValValTrrPileThrGlyLeuSer 175
DB 2 TCTGATAGACAGAGTTGCTGGGACAAAAGCTGTGCTGTGATACAGACTCAGT 61
QY 176 GlySerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHis 195
DB 62 GGTTCAGGAAAAGTACTCTTCTGTGCACTGATCATGAGTTGCATTATAGAGCCAC 121
QY 196 LeuThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPhe 215
DB 122 ATCAGCTATGATCTTGATGGTGCACAACTTAGACATGGCCTCAATCGAGATTTAAGCTTT 181
QY 216 LysAlaGluAspArgAlaGluAsnLeuArgGValGlyGluValAlaLysLeuPheAla 235
DB 182 AAGGAGAAAGACCGTGCAGAAATATACGAAGATTTGTGAAGTGGCAAGCTTTTGGC 241
QY 236 AspAlaGlyValLeuCysIleAlaSerLeuIleSerProTyrArgAspArgAspAla 255
DB 242 GATGCTGGTATCATATGCAATTCAGTTGATATCTCCATACAGAGAGATCGTGATGCA 301
QY 256 CysArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLys 275
DB 302 TGCCGTGCTCTACTTCCAGATTCTAACTTTATTTGAAGTATTTATGGATTTGCCCTTAAA 361
QY 276 IleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLys 295
DB 362 ATTGTGAAGCTCGTATCCTAAAGCTCTGTAAGACTTGCACGACAGAAAGATTAAA 421
QY 296 GlyPheThrGlyIleAspAspProTyrGluProPheIleAsnGlyGluIleValIleLys 315
DB 422 GGTTCCTCAGTGAATTTGATGATCCATACGAAACGCCAGTTAATGTTGAGATGATTAATACG 481
QY 316 MetLysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeu 335
DB 482 ATGAAGAAGTGGGGAATGGCCCTTCAACCAAGCAATGGCGAAGCAAGTCTTATCATACCTT 541
QY 336 GluGluAenGlyTyrLeuGlnAla 343
DB 542 GAAGAGAACGGATATTTGCAAGCT 565
```

RESULT 8

US-10-425-114-16540
; Sequence 16540, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 16540
; LENGTH: 1028
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3066-014-B1_FLI
US-10-425-114-16540

Alignment Scores:
Pred. No.: 1.95e-78 Length: 1028
Score: 905.50 Matches: 187
Percent Similarity: 75.09% Conservative: 21
Best Local Similarity: 67.51% Mismatches: 57
Query Match: 49.59% Indels: 13
DB: 18 Gaps: 2

US-10-829-432-4 (1-343) x US-10-425-114-16540 (1-1028)

```
QY 71 LeuValIleHisGlyLeuThrProArgSerSerHisSerSerAlaGlyLeuAlaSerAsp 90
DB 11 CTCCTCCGCCCGCGCGCTCCAGCGCGTTTGCATCGCCCGCACATCAAGAGCGG 70
QY 91 SerGlyArgGluGlyGluGlyArgGlyAlaArgThrHisCysHisArgGlyIleGly 110
DB 71 CGCGCTCGCGCGCAGCGCGGAGCA-----CGAGGACGACGA 109
QY 111 ArgTrpValArgArgArgArgAsnGlyAlaAlaProGlyGluAlaProHis----- 128
DB 110 CGACGGTCGCGCGCGCGCGCAGAG-GCCGCTTCTAAACGGGTCAAGCGCGCGCGCGCTC 168
QY 129 -----SerProValLysGluLysProValMetSerAsnIleGlyLysSerThrAsn 145
DB 169 CGACGGATCAACAGCAGCAGCAGCGCGCTGTGTGACCTCGACCGCGGGAATCGACGAAC 228
QY 146 IleLeuTyrHisAsnCysLeuIleGlyGlnSerAspArgGlnLysLeuGlyGlnLys 165
DB 229 ATCTGTGGCATAGTGGCCATCGGCAGAGAGCAGCAGCGGTCTGCTGAACCAAGAG 288
QY 166 GlyCysValValTrrPileThrGlyLeuSerGlySerGlyLysSerThrLeuAlaCysAla 185
DB 289 GGCTGCGTGTGTGGATCACTGGCTTAAGCGGTTCAAGGAAAGCACTCGCGTGC 348
QY 186 LeuSerArgGluLeuHisCysArgGlyHisIleThrTyrValLeuAspGlyAspAsnLeu 205
DB 349 CTGACCGCGAGCTGCACGCGAGAGGCCACCTCACGTACGTCCTCGACGGGACAACTC 408
QY 206 ArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAspArgAlaGluAsnIleArg 225
DB 409 AGGCACGGCTGAACAGGGAACCTCAGCTTCGGAGCAGAGGACCGCGCGCAGACATCCGC 468
QY 226 ArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIleAlaSerLeu 245
DB 469 AGAGTAGGGAAGTAGCGAAGCTGTTCCGCCACGCTGGCCTCGTCTGCATCGCCAGCCTC 528
QY 246 IleSerProTyrArgArgAspAlaCysArgAlaLeuLeuProHisSerAsnPhe 265
DB 529 ATATCGCCCTACAGAAGCGACCGAGCGGTGTCCGATCTGCTGCCCAAGCACTCGTTT 588
QY 266 IleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspProLysGlyLeu 285
DB 589 ATCGAGGTGTTCTTCTGGACGTCCGCTTCAAGTGTGCGAAGCCAGGAGCCCAAGGCCTC 648
QY 286 TyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAspProTyrGlu 305
```

```
Db 649 TACAGCTCGCAGCGCCGCAAAATCAAAGGTTTCCACGGCATCGACCATCTTACGAA 708
Qy 306 ProPheAsnGlyGluLeuValLeuLysMetLysAspGluGluCysProSerProLys 325
Db 709 CGCGCGTCGGAGTGTGAGATAGTATCCAGTGTAAAGTCGGCGACTGCGCTTGCCTGAA 768
Qy 326 AlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGlyTyrLeuGln 342
Db 769 TCGATGGCTGGTCACGTTGTGTGCTGACCTTGGAGCAATGGTTCTCTCCAG 819

RESULT 9
US-10-425-115-62889
; Sequence 62889, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 62889
; LENGTH: 2611
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(2611)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MPT4577_157357C.1
US-10-425-115-62889

Alignment Scores:
Pred. No.: 4,528-77 Length: 2611
Score: 896.50 Matches: 196
Percent Similarity: 63.54% Conservative: 34
Best Local Similarity: 54.14% Mismatches: 79
Query Match: 49.10% Indels: 54
DB: 20 Gaps: 7

US-10-829-432-4 (1-343) x US-10-425-115-62889 (1-2611)
Qy 7 AsnGlnThrGluProLeuValThrHisThrGlnProPro----- 20
Db 1151 AGCGCACACGCGCCGCGCACTGCCCATGGACGACGCGCTCCCGGGGCTCTCGTGTCCCTG 1210
Qy 21 -----SerProAlaProGlyProLysSerGlnGlnArgGlnGlnGlnThrLeu 38
Db 1211 TCCGCTCCCGTTGCGGCTCCACCGCAACGCGCGCGCGGGGTGGCGGTCTCTCGCG 1270
Qy 39 -----SerProThrProThrLeuAlaValLeuVal 49
Db 1271 ATCTCGTGTGATTCACGCGCGGGGGTGCCATTCCTCCCTCCCGTGGCGGCTCCCGCGCC 1330
Qy 50 AsnProGlnArg-----AlaProValLeuProGlyLeuThrProSerAsp 65
Db 1331 TCGCGGTCCGCGCACCGGACCGGACCGCGCGCGGCAATCTA-GGGTGGCTCCGCGCTCGC 1389
Qy 66 AlaProLeuProAlaValLeuHisGlyLeuThrProArgSerSerHisSerAla 85
Db 1390 CCCCTCTCCCGAGACTCTAGCGGCGAGGAATTAGGACGAGGTA----- 1434
Qy 86 GlyLeuAlaSerAspSerGlyArgGluGlyGluGlyArg----- 99
Db 1435 -----ACAGCGCACAGAGGCTCGAAGGACGCGCCCTCGTCAATGCGCC 1479
Qy 100 GlyAlaArgThrHisCysHisArgGlyLeuGlyArgTyrValArgArgArgAsn 119
```

```
Db 1480 GCGCGCGCGATCC-----GTGGAGGAGCAGCGCGGAGCAC 1512
Qy 120 GlyAlaAlaProGlyGluAlaProHisSerProValLysGluLysProValMetSerAsn 139
Db 1513 GCGAGGAGTTGAAGGGCAAGCT-----TTGAAAATGTCTATCCACT 1551
Qy 140 IleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGln 159
Db 1552 GTGCCGAGTCAATCAATATCTTCTGGCATGATTCCTTGGCAAGACTGATGGCAAGACTGATGCCAG 1611
Qy 160 LysLeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLys 179
Db 1612 AAGCTACTCAACAGAAAGGCTGTGTCTGTGATCACAGGCTTAGTGGCTCAGGTA 1671
Qy 180 SerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTrpVal 199
Db 1672 AGTACCTTGGCATGTACATTAGGCGCGGAGCTCCACACAGAGGGAAGCTTGCATATGTT 1731
Qy 200 LeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLysAlaGluAsp 219
Db 1732 CTTGACGCTGATTAAGACATGCTCTGAACAGGATCTTGGCTTTAAAGCTGAGAC 1791
Qy 220 ArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyVal 239
Db 1792 CGTCTGAAATATAAGGAGAGATTGGTCAAGTAGCAAGATTATTTCAGATCGCGGCTT 1851
Qy 240 IleCysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeu 259
Db 1852 GTATGTATTGCAAGTTTGTATATCTCATATAGGAGAGACCGTGAATCTTCCGCTGCAATTG 1911
Qy 260 LeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAla 279
Db 1912 TTGCTGTATAGTCTTCAATGAGTTTCTTGAACATGCTCTTGAACATATATGTGAAGA 1971
Qy 280 ArgAspProLysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGly 299
Db 1972 AGAGATCAAGGGCTCTATAAGCTTCTGTGCGAGGAAAAATAAGGGTTTACAGCA 2031
Qy 300 IleAspAspProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGlu 319
Db 2032 ATAGATGACCTTATGAAGCACCACTGAATTTGTGAGATTGAGATCAAGAGGTAGATGTT 2091
Qy 320 GluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluGluAsnGly 339
Db 2092 GTATGCCCTCCGCTGCTGAGATGCGAGGCAAGTTTGTACTTACCTTGAGGAGGAAGGC 2151
Qy 340 TyrLeu 341
Db 2152 TTCCTG 2157

RESULT 10
US-10-425-114-18247
; Sequence 18247, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 18247
; LENGTH: 1180
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
```

Qy	121	AlaAlaProGlyGluAlaProHisSerProVallysGluLysProValMetSerAsnIle	140
Db	6	GCGC	65
Qy	141	GlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlyGlnSerAspArgGlnLys	160
Db	66	GGGAAATCGAGCAACATCTCTGTGGCATGAGTGCGCCATCGGGCAGAAAGAGCGACAGGGT	125
Qy	161	LeuLeuGlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlyLysSer	180
Db	126	CTGCTGAACCAAGAGGGCTGCTGTGTGGATCACTGGCCCTAAGCGGTTACGGGAAAGC	185
Qy	181	ThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeu	200
Db	186	ACGCTCGCGTCGCGCTGAGCCGCGAGCTGCACGGCAGAGGGCCACTCACGTACGTCTCTC	245
Qy	201	AspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheIysAlaGluAspArg	220
Db	246	GACGGCGACACCTCAGGCACCGGGCTGAACAGGGACCTCAGCTTCGGAGCAGAGGACCGC	305
Qy	221	AlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIle	240
Db	306	GCCGAGAACATCCGCGAGNAGGGGAAGTAGCGAAGCTGTTCCGCCAGCGCTGGCCCTCGTC	365
Qy	241	CysIleAlaSerLeuIleSerProTyrArgArgAspArgAspAlaCysArgAlaLeuLeu	260
Db	366	TGCATCGCAGCCCTCATATGCCCTTACAGAAGCGACCGAAGCGGTTGCGCATCTGCTG	425

224	AA	CAGGGTTAAGTGGTTCAGGAAAAACACACTAGCATGCGGCTTAAGTCGCGAGCTGCA	283
191	sCy	AArgGlyHisLeuThrTyrValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnAr	211
284	CT	CAGAGGTCATCTGACCTACATCTTAGCGGTGACAACTCTAAGGCATGGGTCTAAACCG	343
211	gAsp	LeuSerPheLysAlaGluAspArgAlaGluAsnLeuArgArgValGlyLuuValAl	231
344	AG	ACCTCTGTTTGAAGCAAGGACCGGTGTGAATAATATACGCAGTAGTGAGAAGTAGC	403
231	aLys	LeuPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyrArgAr	251
404	AA	AGCTGTTTGACATGCTGGTCTGATCTGCATTGCTAGCTTGATATCACCCCTACAGAAG	463
251	gAsp	ArgAspAlaCysArgAlaLeuLeuProHisSerAsnPhelIleGluValPheIleAs	271
464	TGA	ACGACGGCGCTTTCGCCCAATTTACTGCACAATCTCTACATTCATCGAGAGTGTCTTTTGA	523
271	pLeu	ProLeuLysIleCysGluAlaArgAspProLysGlyLeuTyrLysLeuAlaArgTh	291
524	TGT	CCCACTTGAAGTTGTGAAGCTAGGATCCAAAGGCTTGATCAAGCTTCCCGTGC	583
291	rGly	LysIleLysGlyPheThrGlyIleAspAspProTyrGluProProIleAsnGlyGl	311
584	AGG	AAAAAATCAAAGGGTTTACTGGAATTGATGATCTCTTATGAAGCACCTTCTGCATCGCA	643
311	uIle	ValIleLysMetLysAspGluGluCysProSerProLysAlaMetAlaLysGlnVa	331
644	GAT	AGTATACATGAGTGCAAAGCTGGTGATGCGCCACGCTTAATCGATGGCTGATCAAGT	703
331	lLeu	CysTyrIleuGluGluAsnGlyTyrLeuGln	342
704	TGT	GCATATCTTGAAGCAATGAGTTCTTACAG	737

3

Qy	291	rGLyLysIleLysGlyPheThrGlyIleAspAspProTyrGluProProIleAsnGlyG	311
Db	584	AGGAAAAATCAAAAGGGTTTACTGAAATTGATGATCTTATGAAGCACCTTCTGACTGC	643
Qy	311	uIleValIleLysMetLysAspGluGluCysProSerProLysAlaMetAlaLysGlnVa	331
Db	644	GATAGTGATACAGTGCAAAAGCTGGTGACTGCGCCACGCGCTAAATCGATGGCTGATCAAGT	703
Qy	331	LLeuCysTyrLeuGluGluAsnGlyTyrLeuGln	342
Db	704	TGTGTCAATCTTGAAGCAATAGATTTCTTACAG	737

	: TITLE
:	TITLE
APPLICANT:	

```

; TITLE OR INVENTION: POLYMERIZABLE POLYMERIZATION OF VINYL MONOMERS ASSOCIATED
; WITH OR WITHOUT ACRYLONITRILE
; FILE NO OF INVENTION: 38-21(53313)B
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 1150

```

; :
LENGTH: 1
TYPE: 1

```

; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700102592_FLI
US-10-425-114-1150

Alignment Scores:
Pred. No.:          2.2e-75      Length:        1312
Score:             875.50       Matches:       197
Percent Similarity: 64.07%     Conservative:   33
Best Local Similarity: 54.87%    Mismatches:    86
Query Match:       47.95%      Indels:        46
DB:                18         Gaps:          6

US-10-829-432-4 (1-343) x US-10-425-114-1150 (1-1312)

Qy      7  AsnGlnThrGluprLeuValThrHisThrGlnGlnPro----- 20
      ::::| | | | |
Db      19 AGCGGCACCAAGGCCCGCAGCTGCCCATGGACGACGCTCCCGGGGCTCTCGTGTCCCTG 78

```

Length: 1312

Matches: 197

Conservative: 33

Mismatches: 86

Indels: 46

Gaps: 6

4-1150 (1-1312)

/2TCT-T) OCTY-4

HistThrGlnGlnPro--

[illegible]

CATGGACGACGCTCCCGG


```
QY 21 -----SerProAlaProGlyProAlaSerGlnGlyGlnArgGlnGlnGlyAsnThrLeuLeu 38
Db 79 TCCGCTCCCGTTTCGGGCTCCACAGCAACGCGCGGGTGGCGGTGCTCTCGG 138
QY 39 -----SerProThrProThrLeuAlaValLeuVal 49
Db 139 ATCTCTGTCGATTCACAGCGCGGGGTGCCATTCCTCCCTCGCGGTTCGCGCGCC 198
QY 50 AnProGlnArg-----AlaProValLeuProGlyLeuThrProSerAsp 65
Db 199 TCGCGTCCGACCCGCAACCGGACCGCGCGCGCAATCGA-GGGTTGCTCCGCTCGC 257
QY 66 AlaProLeuAlaSerHisGlyLeuThrProArgSerSerHisSerAla 85
Db 258 CCCCTCTCCGAGACTCTAGCGCAGGATTAGCAGCGGTA----- 302
QY 86 GlyLeuAlaSerAspSerGlyArgGluGlyGluGlyArgGlyAlaArgThrHisCys 105
Db 303 -----ACAGGCACAGAGGCTCGAAGGACGCC-CTTCGTCAA----- 340
QY 106 HisArgGlyIleGlyArgTrpValArgArgArgAenGlyAlaAlaProGlyGlu 125
Db 341 -----ATGCGCGCGCAGCCGATCCGTGGAGAGCAGCCGAGC- 378
QY 126 AlaProHisSerProValysGlyProValMet-----SerAsnIleGlyLys 142
Db 379 ACGCAGGAGGTAGTCTAGTTCAGGCGCAAGCTTTGAAAATGTCATCCATGTGCGGAAG 438
QY 143 SerThrAsnIleLeuTrpHisasnCysLeuIleGlyGlnSerAspArgGlnLysLeuLeu 162
Db 439 TCATCAAAATATCTCTGGCATGATTCCTAGTTGGCAAGACTGATCGCCAGAAAGTACTC 498
QY 163 GlyGlnLysGlyCysValValTrpIleThrGlyLeuSerGlySerGlySerThrLeu 182
Db 499 AACCAAGAGGCTGTGTCTCGATACAGGCCCTTAGTGCTCAGGTAAGATACCTTG 558
QY 183 AlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeuThrTyrValLeuAspGly 202
Db 559 GCATGTACATTAGCGCGGAGCTCCACACAGAGGGAAGCTTGTCATATGTTCTTGACGGT 618
QY 203 AspAsnLeuArgHisGlyLeuasnArgAspLeuSerPheLysAlaGluAspArgAlaGlu 222
Db 619 GATAACTTTAAGACATGGTCTGAACAGGATCTTGGCTTTAAAGCTGAAGACCGTGTGAA 678
QY 223 AsnIleArgArgValGlyGluValAlaLysLeuPheAlaAspAlaGlyValIleCysIle 242
Db 679 AATATAAGGAGAGTTGGTAGTAGCAAAAGTTATTTTCAGACCGCGGCTGTGTATGATT 738
QY 243 AlaSerLeuIleSerProTyrArgArgAspAlaCysArgAlaLeuLeuProHis 262
Db 739 GCAAGTTTGATATCTCCATATAGGAGACCGTGAATCTTGCCTGCAATTGTTATCGGAT 798
QY 263 SerAsnPheIleGluValPheIleAspLeuProLeuLysIleCysGluAlaArgAspPro 282
Db 799 GGTAGCTTTATTAAGATTTCTCTGAATATGCTCTTGGAAATATGATGAAGCAAGGACCCA 858
QY 283 LysGlyLeuTyrLysLeuAlaArgThrGlyLysIleLysGlyPheThrGlyIleAspAsp 302
Db 859 AAGGGCTCTATAGCTTGTCTGCTGGGAAAATTAAGGGTTTTCAGGAAATAGATGAC 918
QY 303 ProTyrGluProProIleAsnGlyGluIleValIleLysMetLysAspGluCysPro 322
Db 919 CCTTATGAAGCACCACTGAATTTGTGAGATTGAGATCAAGGAATAATAGATGGTGTATGCCCT 978
QY 323 SerProLysAlaMetAlaLysGlnValLeuCysTyrLeuGluLysGlyTyrLeu 341
Db 979 TCGCCTTCTGACATGGCGGGCAAGTGGTTACTTACCTTGAGGAAAAGGCTTCTCTG 1035
```

RESULT 14

US-10-425-115-135127

; Sequence 135127, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

```
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 135127
; LENGTH: 1764
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_54717C.1
US-10-425-115-135127
```

Alignment Scores:

Pred. No.:	3,5e-73	Length:	1764
Score:	854.50	Matches:	203
Percent Similarity:	50.41%	Conservative:	32
Best Local Similarity:	52.19%	Mismatches:	83
Query Match:	46.80%	Indels:	73
DB:	20	Gaps:	10

US-10-829-432-4 (1-343) x US-10-425-115-135127 (1-1764)

```
QY 4 HisPheIleAsnGlnThrGluProLeuValThrHisThrGln----- 17
Db 356 CATTTTGAAGTCAACTTGTCTTTTAACTACTCATATAGAGTTTAGTCACACTTCAAGT 415
QY 18 ----- 19
Db 416 TTTAAACATGTTTATTGACAGTCCGGATGGACGACGCTCCCGGCGGCCACCACCCG 475
QY 20 ProSer-----ProAlaProGlyPro-----AlaSerGln 29
Db 476 CCCGCGTGGCGTCCCTGTCCGCTCCGCTTTCGCGCTCCACCGCAACGCGCGCGC 535
QY 30 Gly-----GlnArgGlnGlyAsnThrLeuLeu 38
Db 536 GGGGTGTGTGTCCTCGCGTCTCTGTGTCGATCCAGCGCGCAGGGTGCTGTGCTT 595
QY 39 Ser-----ProThrProThr-LeuAlaValIleLeuValAsnProGlnArgAlaPr 55
Db 596 TGCTTCCCTGGCGTTCGCCCGCTCGCGTTCGCGCACCTGAAC-----CGGAGCCC 649
QY 55 oProValLeuProGlyLeuThrProSerAspAlaProLeuProAlaLeuValIleHisG1 75
Db 650 CGCGTCAAAATCTAGGTTGTCATCCGCGCGGCCCTCTCTCCGAGACTC----- 698
QY 75 yLeuThrProArgSerSerHisSerSerAlaGlyLeuAlaSerAspSerGlyArgArgG1 95
Db 699 -----CCAGCGCTGGGAAC-CAGAACGAGGGGTGACGCGCAGC 735
QY 95 uGlyGluGlyArgGlyAlaArgThrHisCysHisArgGlyIleGlyArgTrpValArgAr 115
Db 736 AAAGTCGACGGACGGCCCCCTCGTCGNAATGCAC-----TGG-----CG 774
QY 115 gArgArgArgAsnGlyAlaAlaProGlyGluAlaProHisSerProValLysGluLys-- 134
Db 775 ACCGTTCGTTCGAGAGCAGCGCGGAGCAGCGCAGGAGGTT-AGTTTAGTTGAGGAAAAGC 833
QY 135 ----ProValMetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeu11 153
Db 834 TTTGAAATGTCATCCACCCTGCGCAAGTCACTCAAAATATCTCTGGCATGATTGTCCAGT 893
QY 153 eGlyGlnSerAspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrpIleThrG1 173
Db 894 TGGCAAGACTGATCGCCAGAGGCTACTCAACAGAAAGGCTGTGTGTCTGGATTACAGG 953
QY 173 yLeuSerGlySerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysAr 193
```



```
Db 954 CCTTAGTGGTTCAGGTAAGAGTACCTTGGCATGTACATTAGGCGCGTGGAGCTCCACACAAG 1013
QY 193 GGLYHisLeuThrTyValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLe 213
Db 1014 AGGGAAGCTTGCATATGCTTGGCGGTGATTAACCTTAAGACATGGTCTGAACAAGGATCT 1073
QY 213 uSerPheLysAlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLe 233
Db 1074 TGGCTTCAAGCTGAAGACCGTGTCTGAAATATAAGGACAGTTGGTGAAGTAGCAAGTT 1133
QY 233 uPheAlaAspAlaGlyValIleCysIleAlaSerLeuIleSerProTyArgArgAspAr 253
Db 1134 ATTTCAGATGCGGGCTTGTATGATGATTCGAAGTTTGTATATCTCCATATAGGAGACCG 1193
QY 253 gAspAlaCysArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuPr 273
Db 1194 TGAATCTTCCGTCGCAATTTGTGTCTGATAGTACCTTCATGAAAGTTTCTTGAACATGTC 1253
QY 273 tGLeuLysIleCysGluAlaArgAspProLysGlyLeuTyLysLeuAlaArgThrGlyLy 293
Db 1254 CITGGAATATGTGAAGCAAGAGATCCAAAGGCGCTTATAGCTTGTCTGTCGAGAAA 1313
QY 293 sIleLysGlyPheThrGlyIleAspAspProTyGluProProIleAsnGlyGluIleVa 313
Db 1314 AATAAAGGGTTTTACAGGAATAGATGACCCCTTATGAAGCACCACCTGAATTGTGAGATTGA 1373
QY 313 lIleLysMetLysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCy 333
Db 1374 GATCAAGAGGTAGATGGTGTATGCGCTCCGCTCTGCTGAGATGGCGGCAAGTTGTATAC 1433
QY 333 sTyLLeuGluGluAsnGlyTyLLeu 341
Db 1434 TTACCTTGAGGAGAAAGGCTTCTCTG 1458
```

RESULT 15

```
US-10-425-114-9652
; Sequence 9652, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 9652
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700856117_FLI
US-10-425-114-9652
```

Alignment Scores:

Pred. No.:	2,26e-73	Length:	1116
Score:	854.00	Matches:	155
Percent Similarity:	87.92%	Conservative:	27
Best Local Similarity:	74.88%	Mismatches:	25
Query Match:	46.77%	Indels:	0
DB:	18	Gaps:	0

US-10-829-432-4 (1-343) x US-10-425-114-9652 (1-1116)

```
QY 137 MetSerAsnIleGlyLysSerThrAsnIleLeuTrpHisAsnCysLeuIleGlnSer 156
Db 281 ATGGCAACTCTGAGCAATTCACATAATATCTTTTGGCAAGATTGTCAAGTAGGAGGCTT 340
```

```
QY 157 AspArgGlnLysLeuLeuGlyGlnLysGlyCysValValTrrIleThrGlyLeuSerGly 176
Db 341 GAAAGGCGAGAGCTACTTAAACCAAAAGGAGTGTGTGTATGGATTACTGGACTCAGCGGA 400
QY 177 SerGlyLysSerThrLeuAlaCysAlaLeuSerArgGluLeuHisCysArgGlyHisLeu 196
Db 401 TCAGGAAAAAGCACATTGGCATGTTCCCTTAAGCAGAGAACTGCACCTCAAAAGGGAAGTTA 460
QY 197 ThrTyValLeuAspGlyAspAsnLeuArgHisGlyLeuAsnArgAspLeuSerPheLys 216
Db 461 TCTTATGTCCTTGATGGAGATAACCTTCGGCATGGACTTAAACCAAGGATCTTGGTTTCAA 520
QY 217 AlaGluAspArgAlaGluAsnIleArgArgValGlyGluValAlaLysLeuPheAlaAsp 236
Db 521 GCCGAGAGATCGCGCTGAAAATATTCCGAGAACTGGAGAAAGTGGCAAAACTGTTTGTCT 580
QY 237 AlaGlyValIleCysIleAlaSerLeuIleSerProTyArgArgAspAlaCys 256
Db 581 GCTGGCTTAAATATGTGTGTGTAGTCTGATATCTCTTACAGAGAGACCGACACACTTGC 640
QY 257 ArgAlaLeuLeuProHisSerAsnPheIleGluValPheIleAspLeuProLeuLysIle 276
Db 641 CGTGCCATGTTTACCTGATGCAAAATTTTATTGAGGTTTTTCATGAACATGCCTCTAGA 700
QY 277 CysGluAlaArgAspProLysGlyLeuTyLysLeuAlaArgThrGlyLysIleLysGly 296
Db 701 TGTGAGGCGACAGATCCAAAGGCTCTATAGCTTGTCTGCGGGGAAATCAAGGT 760
QY 297 PheThrGlyIleAspAspProTyArgProIleAsnGlyGluIleValIleLysMet 316
Db 761 TTTACCGGCATTTGATGATCCTTATGAACCCACCCATAAACTGTGAGATTGAAATAAGCAG 820
QY 317 LysAspGluGluCysProSerProLysAlaMetAlaLysGlnValLeuCysTyLysLeu 336
Db 821 GAAATGGGATTTGTCCAACACCCCACTTTGATGGCTGGACAAAGTAGTTACTTACTTGGAG 880
QY 337 GluAsnGlyTyLLeuGlnAla 343
Db 881 AATAAAGGATTTCTTGAATCC 901
```

Search completed: September 9, 2005, 15:30:30
Job time : 674 secs

THIS PAGE BLANK (USPTO)